

STIC-Biotech/ChemLib

174119

From: Fredman, Jeffrey
Sent: Wednesday, February 11, 2004 11:07 AM
To: STIC-Biotech/ChemLib
Cc: Chernyshev, Olga
Subject: FW: sequence search request

PLEASE RUSH.

I Approve this large search.

Jeff Fredman

RECEIVED
FEB 11 2004
STIC/BIOTECH/CHEN, OLGA
(STIC)

-----Original Message-----

From: Chernyshev, Olga
Sent: Wednesday, February 11, 2004 10:12 AM
To: Fredman, Jeffrey
Subject: sequence search request

Please authorize the following sequence search request:

Please search US case 09/901,187 SEQ ID NOS: 1, 2, 3, 4, 5, 6, 7, 8, 9,10, 11 and 12 in regular and pending databases. Only direct hits, because they claim these short peptides precisely, no open language.
Thank you very much!

Olga N. Chernyshev
AU1646
REM 4E84
20870
mail box 4D70

4C70

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: 12
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: OSP
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-11
Perfect score: 44
Sequence: 1 THIHPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 20938283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	7	10	US-09-901-187B-11
2	43	97.7	148	15	US-10-187-267A-19
3	39	88.6	178	12	US-09-801-944B-233
4	39	88.6	402	12	US-10-341-434-14
5	37	84.1	262	15	US-10-183-116-99
6	37	84.1	270	15	US-10-183-116-91
7	37	84.1	911	12	US-10-115-482-50
8	37	84.1	915	12	US-10-115-482-48
9	37	84.1	1021	12	US-10-017-161-740
10	37	84.1	1818	12	US-10-335-687A-2
11	37	84.1	1824	12	US-10-335-687A-5
12	36	81.8	46	15	US-10-082-830-180
13	35	79.5	49	9	US-09-864-761-38417
14	35	79.5	51	11	US-09-272-975-52
15	35	79.5	362	12	US-10-369-493-4227

16	35	79.5	570	12	US-10-369-493-3132	Sequence 3132, Ap
17	35	79.5	812	12	US-10-369-493-9134	Sequence 9134, Ap
18	35	79.5	1821	12	US-10-372-683-22	Sequence 22, Appl
19	34	77.3	58	15	US-10-083-357-1179	Sequence 1179, Ap
20	34	77.3	253	15	US-10-102-627-66	Sequence 66, Appl
21	34	77.3	338	10	US-09-847-010-6	Sequence 6, Appli
22	34	77.3	338	15	US-10-235-939-6	Sequence 6, Appli
23	34	77.3	344	9	US-09-012-135A-3	Sequence 3, Appli
24	34	77.3	344	12	US-10-295-027-203	Sequence 203, App
25	34	77.3	344	12	US-10-173-999-101	Sequence 101, App
26	34	77.3	344	15	US-10-059-585-34	Sequence 34, Appl
27	34	77.3	344	15	US-10-171-311-214	Sequence 214, App
28	34	77.3	354	12	US-10-108-260A-3493	Sequence 136, App
29	34	77.3	417	12	US-10-369-493-17203	Sequence 3493, Ap
30	34	77.3	439	9	US-09-741-669-300	Sequence 17203, A
31	34	77.3	439	9	US-09-815-242-10284	Sequence 300, App
32	34	77.3	461	9	US-09-731-557A-2	Sequence 10284, A
33	34	77.3	802	11	US-09-946-374-260	Sequence 2, Appli
34	34	77.3	802	12	US-10-015-387A-260	Sequence 260, App
35	34	77.3	802	12	US-10-006-130A-260	Sequence 260, App
36	34	77.3	802	12	US-10-199-672-312	Sequence 312, App
37	34	77.3	802	12	US-10-006-172A-260	Sequence 260, App
38	34	77.3	802	12	US-10-187-749-312	Sequence 312, App
39	34	77.3	802	12	US-10-194-457-312	Sequence 312, App
40	34	77.3	802	12	US-10-184-642-312	Sequence 312, App
41	34	77.3	802	12	US-10-196-747-312	Sequence 312, App
42	34	77.3	802	12	US-10-015-392A-260	Sequence 260, App
43	34	77.3	802	12	US-10-017-253A-260	Sequence 260, App
44	34	77.3	802	12	US-10-173-689-312	Sequence 312, App
45	34	77.3	802	12	US-10-173-689-312	Sequence 312, App

ALIGNMENTS

RESULT 1
US-09-901-187B-11
; Sequence 11, Application US/09901187B
; Patent No. US20020151464A1
; GENERAL INFORMATION:
; APPLICANT: Panacea Pharmaceuticals, Inc.
; APPLICANT: Wolozin, Benjamin
; APPLICANT: Ostretova-Golts, Natalie
; APPLICANT: Lebowitz, Micheal S.
; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatment
; TITLE OF INVENTION: Alpha-Synuclein Diseases
; FILE REFERENCE: PAN01/002US
; CURRENT APPLICATION NUMBER: US/09/901,187B
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/217,319
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/279,199
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent version 3.1
; SEQ ID NO 11
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-187B-11

Query Match 100.0%; Score 44; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05; 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 1 THIHPS 7

RESULT 2
US-10-187-267A-19
; Sequence 19, Application US/10187267A


```

; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286,068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286,292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/287,213
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/288,257
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/291,134
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/291,725
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/294,771
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/296,965
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/299,128
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 149
; SEQ ID NO 50
; LENGTH: 911
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-482-50

Query Match      84.1%; Score 37; DB 12; Length 911;
Best Local Similarity 83.3%; Pred. No. 5.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
DB      320 HLHPS 325
        !:!!!!

RESULT 8
US-10-115-482-48
; Sequence 48, Application US/10115482
; Publication No. US20030212257A1
; GENERAL INFORMATION:
; APPLICANT: Spytex, et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS
; FILE REFERENCE: 21404-322D
; CURRENT APPLICATION NUMBER: US/10/115,482
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/285,325
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 60/285,890
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286,068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286,292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/287,213
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/288,257
; PRIOR FILING DATE: 2001-05-02

```

```
; PRIOR APPLICATION NUMBER: 60/291,134
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/291,725
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/294,771
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/296,965
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/299,128
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 149
; SEQ ID NO 48
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-482-48

Query Match      84.1%; Score 37; DB 12; Length 915;
Best Local Similarity 83.3%; Pred. No. 5.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
Db      320 HLHPS 325

RESULT 9
US-10-017-161-740
; Sequence 740, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 740
; LENGTH: 1021
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-740

Query Match      84.1%; Score 37; DB 12; Length 1021;
Best Local Similarity 83.3%; Pred. No. 6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
Db      361 HLHPS 366

RESULT 10
US-10-335-687A-2
; Sequence 2, Application US/10335687A
; Publication No. US20030166222A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 39267, Human Kinase Family Members and
; TITLE OF INVENTION: Uses Therefor
; FILE REFERENCE: MPI02-001P1RNM
; CURRENT APPLICATION NUMBER: US/10/335,687A
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: 60/345,773
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 149
; SEQ ID NO 48
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-482-48

Query Match      84.1%; Score 37; DB 12; Length 915;
Best Local Similarity 83.3%; Pred. No. 5.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
Db      320 HLHPS 325

RESULT 9
US-10-017-161-740
; Sequence 740, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 740
; LENGTH: 1021
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-740

Query Match      84.1%; Score 37; DB 12; Length 1021;
Best Local Similarity 83.3%; Pred. No. 6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
Db      361 HLHPS 366

RESULT 10
US-10-335-687A-2
; Sequence 2, Application US/10335687A
; Publication No. US20030166222A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 39267, Human Kinase Family Members and
; TITLE OF INVENTION: Uses Therefor
; FILE REFERENCE: MPI02-001P1RNM
; CURRENT APPLICATION NUMBER: US/10/335,687A
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: 60/345,773
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 149
; SEQ ID NO 48
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-482-48

Query Match      84.1%; Score 37; DB 12; Length 1818;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
Db      1223 HLHPS 1228

RESULT 11
US-10-335-687A-5
; Sequence 5, Application US/10335687A
; Publication No. US20030168222A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 39267, Human Kinase Family Members and
; TITLE OF INVENTION: Uses Therefor
; FILE REFERENCE: MPI02-001P1RNM
; CURRENT APPLICATION NUMBER: US/10/335,687A
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: 60/345,773
; PRIOR FILING DATE: 2002-01-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1824
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-335-687A-5

Query Match      84.1%; Score 37; DB 12; Length 1824;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 HHHPS 7
Db      1223 HLHPS 1228

RESULT 12
US-10-082-830-180
; Sequence 180, Application US/10082830
; Publication No. US20030077604A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Salceda, Susana
; APPLICANT: Liu, Chenguua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; FILE REFERENCE: DEX-0249
; CURRENT APPLICATION NUMBER: US/10/082,830
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,802
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 180
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-082-830-180
```

Query Match 81.8%; Score 36; DB 15; Length 46;
 Best Local Similarity 83.3%; Pred. No. 48;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 THHHP 6
 Db 27 SHHHP 32

RESULT 13
 US-09-864-761-38417
 ; Sequence 38417, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aemica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 38417
 ; LENGTH: 49
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO Z98884.11
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.67
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.66

Query Match 79.5%; Score 35; DB 9; Length 49;
 Best Local Similarity 66.7%; Pred. No. 73;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 7
 Db 22 HVHHP 27

RESULT 14
 US-09-272-975-52
 ; Sequence 52, Application US/09272975
 ; Publication No. US2003002774A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hendrickson, Ronald C.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Houghton, Raymond L.
 ; TITLE OF INVENTION: TUBERCULOSIS ANTIGENS AND METHODS
 ; FILE REFERENCE: 210121.474
 ; CURRENT APPLICATION NUMBER: US/09/272,975
 ; CURRENT FILING DATE: 1999-03-18
 ; NUMBER OF SEQ ID NOS: 63
 ; SOFTWARE: PastSeq for Windows Version 4.0
 ; SEQ ID NO 52
 ; LENGTH: 51
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 ; US-09-272-975-52

Query Match 79.5%; Score 35; DB 11; Length 51;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 6
 Db 5 HHHP 9

RESULT 15
 US-10-369-493-4227
 ; Sequence 4227, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xiaofeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 4227
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: Neurospora crassa
 ; US-10-369-493-4227

Query Match 79.5%; Score 35; DB 12; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 6

Query Match 81.8%; Score 36; DB 15; Length 46;
 Best Local Similarity 83.3%; Pred. No. 48;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 THHHP 6
 Db 27 SHHHP 32

RESULT 13
 US-09-864-761-38417
 ; Sequence 38417, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aemica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 38417
 ; LENGTH: 49
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO Z98884.11
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.67
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.66

Query Match 79.5%; Score 35; DB 9; Length 49;
 Best Local Similarity 66.7%; Pred. No. 73;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 7
 Db 22 HVHHP 27

RESULT 14
 US-09-272-975-52
 ; Sequence 52, Application US/09272975
 ; Publication No. US2003002774A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hendrickson, Ronald C.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Houghton, Raymond L.
 ; TITLE OF INVENTION: TUBERCULOSIS ANTIGENS AND METHODS
 ; FILE REFERENCE: 210121.474
 ; CURRENT APPLICATION NUMBER: US/09/272,975
 ; CURRENT FILING DATE: 1999-03-18
 ; NUMBER OF SEQ ID NOS: 63
 ; SOFTWARE: PastSeq for Windows Version 4.0
 ; SEQ ID NO 52
 ; LENGTH: 51
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 ; US-09-272-975-52

Query Match 79.5%; Score 35; DB 11; Length 51;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 6
 Db 5 HHHP 9

RESULT 15
 US-10-369-493-4227
 ; Sequence 4227, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xiaofeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 4227
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: Neurospora crassa
 ; US-10-369-493-4227

Query Match 79.5%; Score 35; DB 12; Length 362;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHP 6

Thu Feb 12 09:07:31 2004

Db |||||
 24 HIHP 28

Search completed: February 11, 2004, 17:54:13
Job time : 24.6667 secs

us 03 001 1070 11111111

GenCore version 5.1.6
Copyright (C) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-901-187C-11
Perfect score: 44
Sequence: 1 THIHPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.76.*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	88.6	312	D86488	bacterial IS-eleme
2	39	88.6	402	I67760	transposase - Esch
3	39	88.6	402	A50737	probable transposa
4	37	84.1	66	C82702	hypothetical prote
5	37	84.1	69	C81075	hypothetical prote
6	37	84.1	611	T22855	hypothetical prote
7	37	84.1	722	S57246	ventral nervous sy
8	36	81.8	156	G31900	hypothetical prote
9	36	81.8	376	S45107	hypothetical prote
10	36	81.8	443	T39608	phage virion transc
11	36	81.8	827	T39608	zinc finger transc
12	35	79.5	169	A82736	hypothetical prote
13	35	79.5	250	C83397	pyrroloquinoline q
14	35	79.5	251	S20455	pgc protein - Kle
15	35	79.5	617	A56051	myocyte nuclear fa
16	35	79.5	693	1 TMBEP7	73.8k alpha trans-
17	35	79.5	1234	A34911	band 3-related pro
18	35	79.5	1237	A31789	band 3-related pro
19	35	79.5	1237	A56764	band 3-related pro
20	35	79.5	1240	S21086	anion exchange pro
21	35	79.5	1530	S23239	brefeldin a resist
22	35	79.5	1530	T52010	hypothetical prote
23	35	79.5	1820	A55494	latent transformin
24	34	77.3	108	A41986	hypothetical prote
25	34	77.3	338	B64148	hypothetical prote
26	34	77.3	392	T139521	rubredoxin-NAD+ re
27	34	77.3	417	E83806	homoserine dehydro
28	34	77.3	434	A23448	guanine deaminase
29	34	77.3	439	C65072	Guanine deaminase

30 34 77.3 439 2 D91098 hypothetical prote
31 34 77.3 439 2 H85943 hypothetical prote
32 34 77.3 474 2 S66480 carbon catabolite
33 34 77.3 515 2 H86202 hypothetical prote
34 34 77.3 520 2 T30808 hypothetical prote
35 34 77.3 566 2 A84485 hypothetical prote
36 34 77.3 704 2 S21911 BCore-NS-Z3 prote
37 34 77.3 774 2 JC6095 hepatocyte nuclear
38 34 77.3 819 2 T10355 hypothetical prote
39 34 77.3 916 2 JT0396 reverse transcript
40 34 77.3 920 2 JC7313 aryl hydrocarbon r
41 34 77.3 1165 2 S27809 GPhase-activating
42 33 75.0 63 2 D82841 hypothetical prote
43 33 75.0 65 2 T18064 hypothetical prote
44 33 75.0 75 2 S25719 protein ECL - rat
45 33 75.0 250 2 T48684 hypothetical prote

ALIGNMENTS

RESULT 1

D86488
bacterial IS-element [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002
C:Accession: D86488
R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86488
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-312 <STO>
A:Cross-references: GB:AE005172; NID:gl1094742; PIDN:AA29676.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: Escherichia coli insertion sequence IS10 transposase

Query Match 88.6%; Score 39; DB 2; Length 312;
Best Local Similarity 85.7%; Pred: No. 7.8; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7

DB 157 THCHPS 163

RESULT 2

I67760
transposase - Escherichia coli insertion sequence IS10
C:Species: Escherichia coli
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 04-Mar-2000
C:Accession: I67760
R:Boysian, G.; Bilyeu, K.; O'Neil, J.P.
Gene 133, 17-22, 1993
A:Title: Genome rearrangements by residual IS10 elements in strains of Escherichia coli
A:Reference number: I53672; MUID:94040791; PMID:8224890
A:Accession: I67760
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-402 <RES>
A:Cross-references: GB:S67119; NID:g455674; PIDN:AA28848.1; PID:g455676
A:Experimental source: plasmid pXT107; insertion sequence IS10
C:Genetics:

A:Mobile element: insertion sequence IS10
C:Superfamily: Escherichia coli insertion sequence IS10 transposase
C:Keywords: DNA binding

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THHHPS 7
|||
Db 247 THCHPS 253

RESULT 3

AE0737
probable transposase STY2056 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AE0737
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, E.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moulle, S.; O'Gaora, P.
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AE0737
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-402 <FAR>
A:Cross-references: GB:AL513382; PIDN:CAD05600.1; PID:gl6503097; GSPDB:GN00176
C:Genetics:
A:Gene: STY2056
C:Superfamily: Escherichia coli insertion sequence IS10 transposase

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THHHPS 7
|||
Db 247 THCHPS 253

RESULT 4

C82702
hypotheoretical protein XFI277 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: C82702
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: AB2315; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: C82702
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-66 <SIM>
A:Cross-references: GB:AE003961; GB:AE003849; NID:g9106254; PIDN:AAF84086.1; GSPDB:GN00176
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.; Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira

M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XFI277

Query Match 84.1%; Score 37; DB 2; Length 66;
Best Local Similarity 85.7%; Pred. No. 3.1;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THHHPS 7
|||
Db 41 THHHPS 47

RESULT 5

C81075
hypotheoretical protein NMB1503 [imported] - Neisseria meningitidis (strain MCS8 serogroup C)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: C81075
R:Rettelin, H.; Saunders, N.J.; Heideberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J. Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizzi, M. Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MCS8.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: C81075
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-69 <TET>
A:Cross-references: GB:AE002500; GB:AE002098; NID:g7226744; PIDN:AAF41859.1; PID:g72267 A:Experimental source: serogroup B, strain MCS8
C:Genetics:
A:Gene: NMB1503

Query Match 84.1%; Score 37; DB 2; Length 69;
Best Local Similarity 83.3%; Pred. No. 3.3;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHPHS 7
|||
Db 40 HHHPHS 45

RESULT 6

T22855
hypotheoretical protein F57F5.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22855
R:Harris, B.
submitted to the EMBL Data Library, July 1996
A:Reference number: Z19627
A:Accession: T22855
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-611 <WIL>
A:Cross-references: EMBL:Z75953; PIDN:CAB00100.1; GSPDB:GN00023; CESP:F57F5.4
C:Genetics:
A:Gene: CESP:F57F5.4
A:Map position: 5
A:Introns: 26/3; 56/2; 101/1; 150/3; 296/3; 382/3; 486/3; 570/1

Query Match 84.1%; Score 37; DB 2; Length 611;
Best Local Similarity 83.3%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HHHPHS 7
|||
Db 451 HHHPHS 456

RESULT 7

S57246
ventral nervous system defective protein - fruit fly (Drosophila melanogaster)
N/Alternate names: NK-2 homeotic protein
C/Species: Drosophila melanogaster
C/Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 15-Oct-1999
C/Accession: S57246; B33976
R/Jimenez, F.; Martin-Morris, L.E.; Velasco, L.; Chu, H.; Sierra, J.; Rosen, D.R.; White
EMBO J. 14, 3487-3495, 1995
A/Title: vnd, a gene required for early neurogenesis of Drosophila, encodes a homeodomain
A/Reference number: S57246; MUID:95354667; PMID:7628450
A/Accession: S57246
A/Molecule type: mRNA
A/Residues: 1-722 <JIM>
A/Cross-references: EMBL:X87141; NID:gl045047; PIDN:CAA60619.1; PID:gl045048
R/Kim, Y.; Nirenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 86, 7716-7720, 1989
A/Title: Drosophila NK-homeobox genes.
A/Reference number: A33976; MUID:90046666; PMID:2573058
A/Accession: B33976
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 476-631, 'VG' <KIM>
A/Cross-references: GB:M27290; NID:gl57635; PIDN:AAA28617.1; PID:gl57636
C/Genetics:
A/Gene: FlyBase:vnd
A/Cross-references: FlyBase:FBgn0003986
C/Superfamily: unassigned homeobox proteins; homeobox homology
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
P:545-601/Domain: homeobox homology <HGX>

Query Match

Best Local Similarity 84.1%; Score 37; DB 2; Length 722;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7

Db 620 THHPHS 626

RESULT 8

GB1900
hypothetical protein NMA1319 [imported] - Neisseria meningitidis (strain Z2491 serogroup
C/Species: Neisseria meningitidis
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C/Accession: GB1900
R/Parthill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagals, K.; Leather, S.; Moulé, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A/Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A/Reference number: A81775; MUID:20222556; PMID:10761919
A/Accession: GB1900
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-156 <PAR>
A/Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAE84570.1; PID:g737999
A/Experimental source: serogroup A, strain Z2491
C/Genetics:
A/Gene: NMA1318; NMA1319

Query Match

Best Local Similarity 81.8%; Score 36; DB 2; Length 156;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 THHPHS 7

Db 138 THHPA 143

RESULT 9

S45107

hypothetical protein 1 - Erwinia carotovora

C/Species: Erwinia carotovora
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
C/Accession: S45107
R/Golby, P.; Jones, S.E.; Stephens, S.; Reeves, P.J.; Bycroft, B.; Stewart, G.; William
submitted to the EMBL Data Library, May 1994
A/Description: Global regulation of Erwinia carotovora exoenzyme virulence factors: mul
A/Reference number: S45107
A/Accession: S45107
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-376 <GOL>
A/Cross-references: EMBL:X79474; NID:g496597; PIDN:CAA55982.1; PID:g496598

Query Match

Best Local Similarity 81.8%; Score 36; DB 2; Length 376;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 THIHPS 6

Db 283 THHPHS 288

RESULT 10

GB1122
phage virion protein, probable NMB1109 [imported] - Neisseria meningitidis (strain MC58
C/Species: Neisseria meningitidis
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C/Accession: GB1122
R/Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.
; H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A/Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A/Reference number: A81000; MUID:20175755; PMID:10710307
A/Accession: GB1122
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-443 <TET>
A/Cross-references: GB:AE002460; GB:AE002098; NID:g7226335; PIDN:AAF41500.1; PID:g72263
A/Experimental source: serogroup B, strain MC58
C/Genetics:
A/Gene: NMB1109

Query Match

Best Local Similarity 81.8%; Score 36; DB 2; Length 443;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 THHPHS 7

Db 425 THHPA 430

RESULT 11

T39608
zinc finger transcription factor - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000
C/Accession: T39608
R/Lyons, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
submitted to the EMBL Data Library, May 1998
A/Reference number: Z21866
A/Accession: T39608
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-827 <LYN>
A/Cross-references: EMBL:AL023554; PIDN:CAAL9035.1; GSPDB:GN00067; SPDB:SPBC16G5.16
A/Experimental source: strain 972h-; cosmid c16G5
C/Genetics:
A/Gene: SPDB:SPBC16G5.16
A/Map position: 2
A/Introns: 30/3; 41/1

C:Superfamily: GAL4 zinc binuclear cluster homology
F:11-47/Domain: GAL4 zinc binuclear cluster homology <GL4>

Query Match 81.8%; Score 36; DB 2; Length 827;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHHP 7
|||||
Db 221 HIHHP 226

RESULT 12

AB2736
hypothetical protein XF1010 [imported] - Xylella fastidiosa (strain 945c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: AB2736
R:Anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: AB2545; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number AB2545 below
A:Accession: AB2736
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-169 <SIM>
A:Cross-references: GB:AE003938; GB:AE003849; NID:g9105935; PIDN:AAF83820.1; GSPDB:GN001
A:Experimental source: strain 945c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Doroty, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laig Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, E.C.; Miyaki, C.Y.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1010
Query Match 79.5%; Score 35; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HIHHP 6
|||||
Db 158 HIHHP 162

RESULT 13
C83397
pyrroloquinoline quinone biosynthesis protein C PA1987 [imported] - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83397
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam, ; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: C83397
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-250 <STO>
A:Cross-references: GB:AE004625; GB:AE004091; NID:g9947983; PIDN:AG05375.1; GSPDB:GN001
A:Experimental source: strain PA01

C:Genetics:
A:Gene: pqqC; PA1987

Query Match 79.5%; Score 35; DB 2; Length 250;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHHP 6
|||||
Db 23 HIHHP 27

RESULT 14

S20455
pqqC protein - Klebsiella pneumoniae
C:Species: Klebsiella pneumoniae
C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 08-Oct-1999
C:Accession: S20455; S21840
R:Meulenbergh, J.J.M.; Sellink, E.; Riegman, N.H.; Postma, P.W.
Mol. Gen. Genet. 232, 284-294, 1992
A>Title: Nucleotide sequence and structure of the Klebsiella pneumoniae pqq operon.
A:Reference number: S20452; MUID:92212293; PMID:1313537
A:Accession: S20455
A:Molecule type: DNA
A:Residues: 1-251 <MEU>
A:Cross-references: EMBL:X58778; NID:g43903; PIDN:CAA41581.1; PID:g43907
C:Genetics:
A:Gene: pqqC
Query Match 79.5%; Score 35; DB 2; Length 251;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HIHHP 6
|||||
Db 24 HIHHP 28

RESULT 15
A56051
myocyte nuclear factor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 24-Sep-1999
C:Accession: A56051
R:Bassey-Duby, R.; Hernandez, M.D.; Yang, Q.; Rochelle, J.M.; Seldin, M.F.; Williams, R.
Mol. Cell. Biol. 14, 4596-4605, 1994
A>Title: Myocyte nuclear factor, a novel winged-helix transcription factor under both d A:Reference number: A56051; MUID:94277065; PMID:8007964
A:Accession: A56051
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-617 <BAS>
A:Cross-references: GB:L26507; NID:g508527; PIDN:AAA37529.1; PID:g508528
C:Genetics:
A:Gene: MNF
C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
C:Keywords: DNA binding; transcription factor
F:289-380/Domain: fork head DNA-binding domain homology <FHD>
Query Match 79.5%; Score 35; DB 2; Length 617;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 THIHPS 7
|||||
Db 583 THSHPT 589

Search completed: February 11, 2004, 17:11:54
Job time: 11.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
(without alignments)
63.714 Million cell updates/sec

Title: US-09-901-187C-11

Perfect score: 44

Sequence: 1 THIHPS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	86.4	1878	1 BAZA_HUMAN	Q9uif9 homo sapien
2	37	84.1	341	1 BRN3_CHICK	Q91998 gallus gall
3	37	84.1	611	1 ADD2_CABEL	Q20952 caenorhabdi
4	37	84.1	722	1 HMW2_DROME	P22808 drosophila
5	36	81.8	376	1 YQCB_ERWCA	Q47417 erwilia car
6	36	81.8	384	1 VPAP_PRVXA	P36702 pseudorabie
7	36	81.8	637	1 PQOC_KLEPN	P27505 gallus gall
8	35	79.5	251	1 POOC_KLEPN	P27505 kliebiella
9	35	79.5	617	1 FXK1_MOUSE	P42128 mus musculu
10	35	79.5	693	1 UL47_HSV11	P10231 herpes simp
11	35	79.5	1210	1 PER3_HUMAN	P56645 homo sapien
12	35	79.5	1234	1 B3A2_RAT	P23347 rattus norv
13	35	79.5	1237	1 B3A2_MOUSE	P13808 mus musculu
14	35	79.5	1237	1 B3A2_RABIT	P48746 cryctolegus
15	35	79.5	1238	1 B3A2_CAVPO	Q92058 cavia porce
16	35	79.5	1241	1 B3A2_HUMAN	P04920 homo sapien
17	35	79.5	1530	1 BFR1_SCHPO	P41820 schizosacch
18	35	79.5	1850	1 BAZA_MOUSE	Q91y65 mus musculu
19	34	77.3	338	1 YJFK_HAEIN	P44641 haemophilus
20	34	77.3	344	1 STKC_HUMAN	Q96944 homo sapien
21	34	77.3	392	1 RUCB_ACICA	P24154 acinetobact
22	34	77.3	408	1 HNAG_HUMAN	P42541 homo sapien
23	34	77.3	408	1 HNAG_MOUSE	Q9wu06 mus musculu
24	34	77.3	439	1 GUAD_ECOLI	P76641 escherichia
25	34	77.3	474	1 MIG1_KLULA	Q50898 kluyveromyc
26	34	77.3	819	1 Y033_NPVOP	O10336 crygylia pseu
27	34	77.3	916	1 RTJK_DROME	P13228 drosophila
28	34	77.3	1165	1 GAB1_DROME	P48423 drosophila
29	33	75.0	264	1 OTP_DROME	P56672 drosophila
30	33	75.0	297	1 CLP1_MOUSE	Q08291 mus musculu
31	33	75.0	297	1 CLP1_RAT	Q08290 rattus norv
32	33	75.0	337	1 Y020_BUCAP	Q44634 buchnera ap
33	33	75.0	439	1 GUAD_DEIRA	Q9ryx4 deinococcus

34	33	75.0	465	1 NORA_CHLMU	O9plu3 chlamydia m
35	33	75.0	465	1 NORA_CHLTR	O84639 chlamydia t
36	33	75.0	514	1 VNN1_CANFA	Q9tsx8 canis famli
37	33	75.0	860	1 ATSG6_HUMAN	Q9ukp5 homo sapien
38	33	75.0	1275	1 RGA2_SCHPO	Q10164 schizosacch
39	32	72.7	227	1 IFEL_ORYSA	P48599 oryza sativ
40	32	72.7	368	1 VE2_HPV45	P36794 human papil
41	32	72.7	376	1 HMEV_DROME	P06602 drosophila
42	32	72.7	449	1 GAT6_HUMAN	O92908 homo sapien
43	32	72.7	486	1 BAF1_KLULA	P26375 kluyveromyc
44	32	72.7	518	1 VL2_HPV5B	P26540 human papil
45	32	72.7	592	1 THD1_ARATH	Q9ZS86 arabidopsi

ALIGNMENTS

RESULT 1
BAZA_HUMAN
ID BAZA_HUMAN STANDARD; Q96H26; PRT; 1878 AA.
AC Q9uif9; O00536; O15030; Q96H26;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Bromodomain adjacent to zinc finger domain 2A (Transcription
DE termination factor-1 interacting protein 5) (TTF-1 interacting protein
DE 5) (TIPS) (hWAIp3).
GN BAZA OR TIPS OR KIAA0314.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=20130112; PubMed=10662543;
RA Jones M.H., Hamana N., Nezu J.-I., Shimane M.;
RT "A novel family of bromodomain genes.";
RL Genomics 63:40-45(2000).
RN [2]
RP SEQUENCE OF 332-738 FROM N.A.
RC TISSUE=Lung;
RA Jansa P., Grummt I.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 639-1878 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
RN [4]
RP SEQUENCE OF 1035-1878 FROM N.A.
RC TISSUE=Lymph;
RX MEDLINE=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Hellon E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnurch A., Schein J.E., Jones S.O.M., Marra W.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: May play a role in transcriptional regulation
 CC interacting with ISWI. May serve a specific role in maintaining or
 CC altering the chromatin structure of the rDNA locus (By
 CC similarity).
 CC -!- SUBUNIT: Together with ISWI/SNP2h, it forms a complex termed NORC
 CC (nucleolar remodeling complex).
 CC -!- SUBCELLULAR LOCATION: Nuclear. Co-localizes with the basal RNA
 CC polymerase I transcription factor UBF in the nucleolus.
 CC -!- TISSUE SPECIFICITY: Expressed at moderate levels in most tissues
 CC analyzed, including heart, brain, placenta, lung, skeletal muscle,
 CC kidney and pancreas.
 CC -!- SIMILARITY: BELONGS TO THE WAL FAMILY.
 CC -!- SIMILARITY: Contains 4 A.T hook DNA-binding repeats.
 CC -!- SIMILARITY: Contains 1 methyl-binding (MBD) domain.
 CC -!- SIMILARITY: Contains 1 bromodomain.
 CC -!- SIMILARITY: Contains 1 DDT domain.
 CC -!- SIMILARITY: Contains 1 PHD-type zinc finger.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AB032234; BAA89211.1; -;
 CC EMBL; AF000422; AAB60864.1; -;
 CC EMBL; AB002312; BAA20773.1; -;
 CC EMBL; BC008965; AAB08965.1; -;
 CC HSSP; Q92831; 1B91.
 CC Genew; HGNC:962; BAZ2A.
 CC MIM; 605682;
 CC GO; GO:0005731; C:nucleolus organizer complex; NAS.
 CC GO; GO:0003677; F:DNA binding activity; NAS.
 CC GO; GO:0030528; F:transcription regulator activity; NAS.
 CC GO; GO:0006338; P:chromatin modeling; NAS.
 CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 CC InterPro; IPR000637; AT hook.
 CC InterPro; IPR001487; Bromodomain.
 CC InterPro; IPR004022; DDT dom.
 CC InterPro; IPR001739; Methyl-CpG_bind.
 CC InterPro; IPR001965; Znf PHD.
 CC Pfam; PF02178; AT hook; 4.
 CC Pfam; PF00439; bromodomain; 1.
 CC Pfam; PF02791; DDT; 1.
 CC Pfam; PF01429; MBD; 1.
 CC Pfam; PF00828; PHD; 1.
 CC PRINTS; PR00503; BROMODOMAIN.
 CC SMART; SM00384; AT hook; 4.
 CC SMART; SM00297; BROMO; 1.
 CC SMART; SM00571; DDT; 1.
 CC SMART; SM00391; MBD; 1.
 CC SMART; SM00249; PHD; 1.
 CC PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
 CC PROSITE; PS0014; BROMODOMAIN_2; 1.
 CC PROSITE; PS001359; ZF_PHD_1; FALSE_NEG.
 CC PROSITE; PS0016; ZF_PHD_2; 1.
 CC Transcription regulation; Bromodomain; Zinc-finger; Coiled coil;
 CC Transcription regulation; Repeat; DNA-binding.
 CC MIM; 525 577 MBD.
 CC FT DNA_BIND 525 634 A.T HOOK 1.
 CC FT DNA_BIND 622 655 A.T HOOK 2.
 CC FT DOMAIN 643 883 DDT.
 CC FT DNA_BIND 821 1159 A.T HOOK 3.
 CC FT DNA_BIND 1159 1377 A.T HOOK 4.
 CC FT DNA_BIND 1377 1699 PHD-TYPE.
 CC FT ZN_FING 1649 1853 BROMODOMAIN.
 CC FT DOMAIN 1783 1853

FT DOMAIN 633 772 LYS-RICH.
 FT DOMAIN 666 765 COILED COIL (POTENTIAL).
 FT DOMAIN 1185 1250 GLU-RICH.
 FT DOMAIN 1263 1384 PRO-RICH.
 FT DOMAIN 1732 1735 POLY-ARG.
 FT CONFLICT 574 574 L -> V (IN REF. 2).
 FT CONFLICT 700 700 L -> Q (IN REF. 2 AND 3).
 FT CONFLICT 720 720 H -> Q (IN REF. 2 AND 3).
 FT CONFLICT 727 738 SKAEKGGTKQ -> KIKKKKKKKK (IN REF. 2).
 FT CONFLICT 785 785 K -> R (IN REF. 3).
 FT CONFLICT 951 951 P -> L (IN REF. 3).
 FT CONFLICT 1006 1006 GR -> EG (IN REF. 3).
 FT CONFLICT 1037 1037 IAA -> GTR (IN REF. 4).
 FT CONFLICT 1163 1163 G -> S (IN REF. 1).
 FT CONFLICT 1166 1166 R -> L (IN REF. 1).
 FT CONFLICT 1172 1172 S -> F (IN REF. 1).
 FT CONFLICT 1178 1178 A -> V (IN REF. 1).
 FT CONFLICT 1202 1202 L -> F (IN REF. 1).
 FT CONFLICT 1292 1292 P -> L (IN REF. 1).
 FT CONFLICT 1295 1295 L -> F (IN REF. 1).
 FT CONFLICT 1313 1313 P -> L (IN REF. 1).
 FT CONFLICT 1407 1410 MISSING (IN REF. 4).
 FT CONFLICT 1416 1416 R -> P (IN REF. 1).
 FT CONFLICT 1541 1541 R -> P (IN REF. 1).
 FT CONFLICT 1571 1571 E -> K (IN REF. 1).
 FT CONFLICT 1616 1616 V -> I (IN REF. 1).
 FT CONFLICT 1622 1622 E -> Q (IN REF. 1).
 FT CONFLICT 1629 1629 Q -> H (IN REF. 1).
 FT CONFLICT 1636 1636 Q -> H (IN REF. 1).
 FT CONFLICT 1739 1739 R -> K (IN REF. 1).
 FT CONFLICT 1754 1754 G -> R (IN REF. 1).
 FT CONFLICT 1878 1878 AA; 208480 MW; 400970CA68234317 CRC64;
 SQ SEQUENCE
 Query Match 86.4%; Score 38; DB 1; Length 1878;
 Best Local Similarity 71.4%; Pred No. 69;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 THIHPS 7
 Db 83 SHLHPS 89
 ID -BRN3_CHICK STANDARD; PRT; 341 AA.
 AC Q91998;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Brain-specific homeobox/POU domain protein 3 (Brn-3).
 GN BRN3.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Dorsal root ganglion;
 RX MEDLINE=97349928; PubMed=9205808;
 RA Lindeberg J.K., Klint P., Williams R., Ebendal T.;
 RT "Identification of a chicken homologue in the Brn-3 subfamily of POU-
 RT transcription factors";
 RL Brain Res. Dev Brain Res. 100:169-182(1997).
 CC -!- FUNCTION: MAY PLAY A ROLE IN SPECIFYING TERMINALLY DIFFERENTIATED
 CC NEURONAL PHENOTYPES (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: BELONGS TO THE POU TRANSCRIPTION FACTOR FAMILY.
 CC CLASS=4 SUBFAMILY.
 CC -!- SIMILARITY: Contains 1 homeobox domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X87141; CAA50619.1; -
 CC EMBL; M27290; AAA38617.1; -
 CC PIR; S57246; S57246.
 CC PDB; 1VND; 08-NOV-96.
 CC PDB; 1NK2; 23-FEB-99.
 CC PDB; 1NK3; 23-FEB-99.
 CC PDB; 1QRY; 06-JUL-99.
 CC TRANSFAC; T04258; -
 CC FlyBase; FBgn0003986; vnd.
 CC GO; GO:0007400; P:neuroblast cell fate determination; IMP.
 CC GO; GO:0007419; P:ventral cord development; NAS.
 CC InterPro; IPR001356; Homeobox.
 CC Pfam; PF00046; homeobox; 1.
 CC ProDom; PD000010; Homeobox; 1.
 CC SMART; SM00389; HOX; 1.
 CC PROSITE; PS00027; HOMEBOX_1; 1.
 CC PROSITE; PS00071; HOMEBOX_2; 1.
 CC Transcription regulation; Homeobox; DNA-binding; 3D-structure;
 CC Developmental protein; Nuclear protein; Neurogenesis.
 CC DOMAIN 108 111 POLY-SER.
 CC DOMAIN 215 223 POLY-ALA.
 CC DOMAIN 336 343 POLY-ALA.
 CC DOMAIN 351 359 POLY-ALA.
 CC DOMAIN 482 486 POLY-ASN.
 CC DOMAIN 489 492 POLY-ASN.
 CC DOMAIN 544 603 HOMEBOX.
 CC DOMAIN 688 693 POLY-ALA.
 CC VARIANT 578 578 A -> T (IN VND29).
 CC CONFLICT 632 633 RR -> VG (IN REF. 2).
 CC TURN 553 554
 CC HELIX 555 555
 CC HELIX 571 581
 CC TURN 582 582
 CC HELIX 585 594
 CC SEQUENCE 722 AA; 76468 MW; D036AE4D89C0L4DA CRC64;

Query Match 84.1%; Score 37; DB 1; Length 722;
 Best Local Similarity 85.7%; Pred. No. 37;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 THIHPS 7
 |||||
 Db 620 THPHPS 626

RESULT 5
 YQCB_ERWCA STANDARD; PRT; 376 AA.
 AC Q47417;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Exoenzyme regulation regulon ORF1.
 OS Erwinia carotovora.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Pectobacterium.
 CC NCBI_TaxID=554;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=SCRI193;
 CC Golby P., Jones S.E., Stephens S., Reeves P.J., Bycroft B.,
 CC Stewart G., Williams P., Salmund G.P.C.;
 CC "Global regulation of Erwinia carotovora exoenzyme virulence
 CC factors: multicopy suppression of rex mutants and evidence for a
 CC global repression regulon."
 CC Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE RLJ FAMILY OF PSEUDOURIDINE SYNTHASES.
 CC STRONG, TO E.COLI YQCB.
 CC -!- SIMILARITY: IN THE N-TERMINAL, STRONG, TO E.COLI YQCC; IN THE C-
 CC TERMINAL, STRONG, TO E.COLI YQCB.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X79474; CAA55982.1; -
 CC PIR; S45107; S45107.
 CC InterPro; IPR006145; PseudoU synth.
 CC InterPro; IPR006224; PSI_RLJ.
 CC Pfam; PF04287; DUF446; 1.
 CC Pfam; PF06849; PseudoU synth_2; 1.
 CC ProDom; PD001819; PSI_RLJ; 1.
 CC PROSITE; PS01129; PSI_RLJ; 1.
 CC SEQUENCE 376 AA; 43605 MW; F35992CAAAD22E30 CRC64;

Query Match 81.8%; Score 36; DB 1; Length 376;
 Best Local Similarity 83.3%; Pred. No. 28;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 THIHHP 6
 :|||
 Db 283 SHIHHP 288

RESULT 6
 VPAP_PVKA STANDARD; PRT; 384 AA.
 ID VPAP_PVKA
 AC P36702;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE DNA polymerase processivity factor (Polymerase accessory protein)
 DE (PAP) (UL42 homolog).
 OS Pseudorabies virus (strain Kaplan) (PRV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 CC NCBI_TaxID=33703;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=9522727; PubMed=7707503;
 CC Berthomme H., Monahan S.J., Farris D.S., Jacquemont B.,
 CC Epstein A.L.;
 CC "Cloning, sequencing, and functional characterization of the two
 CC subunits of the pseudorabies virus DNA polymerase holoenzyme:
 CC evidence for specificity of interaction."
 CC J. Virol. 69:2811-2818(1995).

CC -!- FUNCTION: ACCESSORY SUBUNIT OF THE DNA POLYMERASE THAT ACTS TO
 CC INCREASE THE PROCESSIVITY OF POLYMERIZATION (BY SIMILARITY).
 CC -!- SIMILARITY: TO OTHER HERPESVIRUSES POLYMERASE ACCESSORY PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M94355; AAA74384.1; -
 CC InterPro; IPR003202; UL42.
 CC Pfam; PF02282; UL42; 2.
 CC DNA-binding; DNA replication.
 CC SEQUENCE 384 AA; 40305 MW; FFAC3EF0C1984936 CRC64;

Query Match 81.8%; Score 36; DB 1; Length 384;

Best Local Similarity 71.4%; Pred. No. 28; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THIHPS 7
16 TAAHPA 22

Db

RESULT 7
SCAA_CHICK STANDARD; PRT; 637 AA.
AC Q92075; P70095; Q98941;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Amiloride-sensitive sodium channel alpha-subunit (Epithelial Na+ channel alpha subunit) (Alpha Enac) (Nonvoltage-gated sodium channel 1 alpha subunit) (SCNEA) (Alpha NaCh).
GN SCNN1A OR ENAC.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; OC Gallus.
OX NCBI_TaxID=9031;
[1]
RN SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RP STRAIN=Isa brown; TISSUE=Cochlea;
RX MEDLINE=97157073; PubMed=9003454;
RA Killick R., Richardson G.;
RT "Isolation of chicken alpha Enac splice variants from a cochlear cDNA library."
RL Biochim. Biophys. Acta 1350:33-37(1997).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Intestine;
RA Goldstein O., Asher C., Garty H.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SODIUM PERMEABLE NON-VOLTAGE-SENSITIVE ION CHANNEL INHIBITED BY THE DIURETIC AMILORIDE. MEDIATES THE ELECTRODIFFUSION OF THE LUMINAL SODIUM (AND WATER, WHICH FOLLOWS OSMOTICALLY) THROUGH THE APICAL MEMBRANE OF EPITHELIAL CELLS. CONTROLS THE REABSORPTION OF SODIUM IN KIDNEY, COLON, LUNG AND SWEAT GLANDS.
CC ALSO PLAYS A ROLE IN TASTE PERCEPTION (BY SIMILARITY).
CC -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA, ONE BETA AND ONE GAMMA SUBUNIT. A DELTA SUBUNIT CAN REPLACE THE ALPHA SUBUNIT (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q92075-1; Sequence=Displayed;
CC Name=Short; 2; Sequence=VSP_006196, VSP_006197;
CC -!- TISSUE SPECIFICITY: THE LONG ISOFORM HAS BEEN FOUND IN COCHLEA, COLON, AND CARTILAGE. THE SHORT ISOFORM IS ONLY FOUND IN COCHLEA.
CC -!- SIMILARITY: BELONGS TO THE AMILORIDE-SENSITIVE SODIUM CHANNEL FAMILY.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-15 IS THE INITIATOR.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; U62902; AAB50550.1; -
CC EMBL; U62903; AAB50551.1; -
CC EMBL; U62904; AAB50552.1; -
CC EMBL; U58475; AAB04954.1; ALT_INIT.
CC InterPro; IPR004724; Enac.
CC InterPro; IPR001873; Na-channel_ASC.

DR Pfam; PF00858; ASC; 1.
DR PRINTS; PR01078; AMINACHANNEL.
DR TIGRFAMs; TIGR00859; ENAC; 1.
DR PROSITE; PS01206; ASC; 1.
KW Ionic channel; Transmembrane; Ion transport; Glycoprotein;
KW Alternative splicing.
FT DOMAIN 1 80
FT TRANSMEM 81 97
FT DOMAIN 98 554
FT TRANSMEM 555 571
FT TRANSMEM 572 637
FT DOMAIN 157 157
FT CARBOHYD 280 280
FT CARBOHYD 298 298
FT CARBOHYD 499 499
FT CARBOHYD 402 434
FT VARSPLIC 435 637
FT VARSPLIC 435 637
FT Missing (in isoform Short).
FT CONFLICT 49 50
FT CONFLICT 578 578
FT CONFLICT Q -> R (IN REF. 2).
FT SEQUENCE 637 AA; 73236 MW; 5D15E616373971DA CRC64;
Query Match 81.8%; Score 36; DB 1; Length 637;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THIHPS 6
282 THFHHP 287

Db

RESULT 8
PQQC_KLEPN STANDARD; PRT; 251 AA.
AC P27505;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Coenzyme PQQ synthesis protein C.
DE PQQ.
GN Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=573;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=NCTC 418;
RX MEDLINE=92212293; PubMed=1313537;
RA Meulenbergh J.J.M., Sellink E., Riegman N.H., Postma P.W.;
RT "Nucleotide sequence and structure of the Klebsiella pneumoniae pqq operon."
RL Mol. Gen. Genet. 232:284-294(1992).
CC -!- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ) BIOSYNTHESIS.
CC -!- SIMILARITY: TO OTHER BACTERIAL PQQC.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; X58778; CAA41581.1; -
CC EMBL; S20455; S20455.
CC Pfam; PF05312; PQQC; 1.
CC PQQ.
CC SEQUENCE 251 AA; 28986 MW; B48A494FA63B6598 CRC64;
CC

Query Match 79.5%; Score 35; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 27; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 2 HIRHP 6
DB 24 HIRHP 28

RESULT 9
FXK1_MOUSE
ID FXK1_MOUSE STANDARD; PRT; 617 AA.
AC P42128; O35939;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Forkhead box protein K1 (Myocyte nuclear factor) (MNF).
GN FOXK1 OR MNF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=94277065; PubMed=8007964;
RA Bassel-Duby R., Hernandez M.D., Yang Q., Rochelle J.M.,
RA Seldin M.F., Williams R.S.;
RT "Transient expression of a winged-helix protein, MNF-beta, during
RT myogenesis";
RT Mol. Cell. Biol. 17:5236-5243 (1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=97415602; PubMed=9271401;
RA Yang Q., Bassel-Duby R., Williams R.S.;
RA "Transient expression of a winged-helix protein, MNF-beta, during
RT myogenesis";
RT Mol. Cell. Biol. 17:5236-5243 (1997).
CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR THAT BINDS TO THE UPSTREAM
CC ENHANCER REGION (CCAC BOX) OF MYOGLOBIN GENE. HAS A ROLE IN
CC MYOGENIC DIFFERENTIATION AND IN REMODELING PROCESSES OF ADULT
CC MUSCLES THAT OCCUR IN RESPONSE TO PHYSIOLOGICAL STIMULI.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P42128-1; Sequence=Displayed;
CC Name=2; Synonyms=Beta;
CC IsoId=P42128-2; Sequence=VSP_001545, VSP_001546;
CC -!- TISSUE SPECIFICITY: EXPRESSED IN TISSUES AND CELLS IN WHICH THE
CC MYOGLOBIN GENE IS TRANSCRIPTIONALLY ACTIVE (CARDIAC AND SKELETAL
CC MYOCYTES, BRAIN, KIDNEY. . .).
CC -!- PTM: PHOSPHORYLATED
CC -!- SIMILARITY: Contains 1 fork-head domain.
CC -!- SIMILARITY: Contains 1 FHA domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; L26507; AAA37529.1; -
DR EMBL; U95016; AAB69641.1; -
DR PIR; A56051; A56051.
DR KSSP; Q83245; 2HPH.
DR TRANSFAC; T04216; -
DR MGD; MGI:1347488; Foxk1.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00498; FHA; 1.

Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR SMART; SM00240; FHA; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS00039; FORK_HEAD_3; 1.
DR PROSITE; PS00006; FHA_DOMAIN; 1.
KW DNA-binding; Transcription regulation; Activator; Nuclear protein;
KW Phosphorylation; Alternative splicing; Differentiation.
FT DOMAIN 27 33 POLY-ALA.
FT DOMAIN 107 159 FHA.
FT DNA_BIND 288 379 FORK-HEAD.
FT VARSPLIC 400 409 RSPASHTPG -> SAPASHTSHA (in isoform 2).
FT VARSPLIC 410 617 /FTID=VSP_001545.
FT VARSPLIC 410 617 Missing (in isoform 2).
FT VARSPLIC 410 617 /FTID=VSP_001546.
SQ SEQUENCE 617 AA; 65839 MW; A1083B28C709FC4A CRC64;

Query Match 79.5%; Score 35; DB 1; Length 617;
Best Local Similarity 71.4%; Pred. No. 69;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
DB 583 THSHPT 589

RESULT 10
UL47_HSV11
ID UL47_HSV11 STANDARD; PRT; 693 AA.
AC P10231;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Virion protein UL47 (82/81 kDa tegument protein) (VIM82/81) (VPI13/14).
GN UL47.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1";
RL J. Gen. Virol. 69:1531-1574 (1988).
RN [2]
RP IDENTIFICATION OF PROTEIN.
RX MEDLINE=91108391; PubMed=2177087;
RA McLean G., Rixon F., Langeland N., Haarr L., Marsden H.;
RT "Identification and characterization of the virion protein products
RT of herpes simplex virus type 1 gene UL47";
RL J. Gen. Virol. 71:2953-2960 (1990).
CC -!- FUNCTION: MODULATOR OF ALPHA-TIF (VIM65 PHOSPHOPROTEIN) TRANS-
CC ACTIVATION. POSSIBLY UL47 MAY HAVE KINASE ACTIVITY.
CC -!- SUBCELLULAR LOCATION: MAJOR TEGUMENT PROTEIN OF THE VIRIONS.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE LATER STAGES OF INFECTION.
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL47,
CC EBV-1 13, AND VZV 11.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; X14112; CAA32297.1; -

FT TRANSMEM 984 1006 POTENTIAL.
FT TRANSMEM 1032 1053 POTENTIAL.
FT TRANSMEM 1087 1132 POTENTIAL.
FT TRANSMEM 1159 1195 POTENTIAL.
FT DOMAIN 5 316 PRO-RICH.
FT DOMAIN 73 87 HIS-RICH.
FT DOMAIN 861 865 POLY-SER.
FT CARBOHYD 855 855 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 866 866 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 878 878 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 1169 1169 PALMITATE (BY SIMILARITY).
FT VARSPLIC 1 17 MSSAPRRPASGADSLHT -> MDFLLRPQ (in isoform B2).
FT VARSPLIC 1 17 /FTID=VSP 000457.
FT VARSPLIC 1 17 MSSAPRRPASGADSLHT -> MTQ (in isoform B1).
FT VARSPLIC 1 166 Missing (in isoform C2).
FT VARSPLIC 1 198 Missing (in isoform C1).
FT VARSPLIC 167 193 /FTID=VSP 000460.
FT VARSPLIC 167 193 ERTSPSPPTQTHQEAAPRASKAQTG -> MPAPFOEWKSG
FT VARSPLIC 167 193 GLREAAVFGAGSVCR (in isoform C2).
FT VARSPLIC 167 193 /FTID=VSP 000461.
FT CONFLICT 205 205 A -> G (IN REF. 2).
FT SEQUENCE 1237 AA; 136813 MW; 1A0782C0071782EE CRC64;
Query Match 79.5%; Score 35; DB 1; Length 1237;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HIHHP 6
Db 79 HIHHP 83
RESULT 14
B3A2 RABIT
ID B3A2 RABIT STANDARD; PRT; 1237 AA.
AC P49736;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Anion exchange protein 2 (Non-erythroid band 3-like protein) (B3RP).
GN SLCA42 OR AE2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9996;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white;
RX MEDLINE=93035730; PubMed=1415547;
RA Chow A., Robbins J.W., Aronson P.S., Igarashi P.;
RT "cDNA cloning and localization of a band 3-related protein from
ileum.";
RL Am. J. Physiol. 263:G345-G352(1992).
CC -!- FUNCTION: PLASMA MEMBRANE ANION EXCHANGE PROTEIN OF WIDE
DISTRIBUTION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE ANION EXCHANGER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; S45791; AAB23488.1;
CC PIR; A56764; A56764.
CC HSSP; P02730; 1BTO.
CC InterPro; IPR001717; Anion_exchange.

DR InterPro; IPR003020; HCO3_cotransp.
DR Pfam; PF00955; HCO3_cotransp; 1.
DR PRINTS; PRO1231; HCO3TRANSPORT.
DR TIGRFAMs; TIGR00834; ae; 1.
DR PROSITE; PS00219; ANION_EXCHANGER_1; 1.
DR PROSITE; PS00220; ANION_EXCHANGER_2; 1.
KW Transmembrane; Glycoprotein; Transport; Anionport; Ion transport;
FT DOMAIN 1 703 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 704 1237 MEMBRANE (ANION EXCHANGE).
FT TRANSMEM 704 727 POTENTIAL.
FT TRANSMEM 733 770 POTENTIAL.
FT TRANSMEM 790 812 POTENTIAL.
FT TRANSMEM 822 843 POTENTIAL.
FT TRANSMEM 844 896 EXOPLASMIC LOOP (POTENTIAL).
FT TRANSMEM 897 914 POTENTIAL.
FT TRANSMEM 915 929 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 930 950 POTENTIAL.
FT TRANSMEM 1032 1053 POTENTIAL.
FT TRANSMEM 1087 1132 POTENTIAL.
FT TRANSMEM 1159 1195 POTENTIAL.
FT TRANSMEM 1159 1195 PRO-RICH.
FT TRANSMEM 5 316 HIS-RICH.
FT TRANSMEM 74 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 855 855 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 864 864 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 878 878 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 1169 1169 PALMITATE (BY SIMILARITY).
FT SEQUENCE 1237 AA; 136535 MW; 2811D11051552B2 CRC64;
Query Match 79.5%; Score 35; DB 1; Length 1237;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HIHHP 6
Db 80 HIHHP 84
RESULT 15
B3A2 CAVPO
ID B3A2 CAVPO STANDARD; PRT; 1238 AA.
AC Q9Z588;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Anion exchange protein 2 (Non-erythroid band 3-like protein) (AE2
DE Anion exchanger).
DE Anion exchanger).
GN SLCA42 OR AE2.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NH 2; TISSUE=Organ of Corti;
RX MEDLINE=99023787; PubMed=9804866;
RA Mnatre A.N., Charachon G., Alper A.L., Lalwani A.K.;
RT "The guinea pig cochlear AE2 anion exchanger: cDNA cloning and in situ
localization within the cochlea.";
RL Biochim. Biophys. Acta 1414:1-15(1998).
CC -!- FUNCTION: PLASMA MEMBRANE ANION EXCHANGE PROTEIN OF WIDE
DISTRIBUTION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE ANION EXCHANGER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; S45791; AAB23488.1;
CC PIR; A56764; A56764.
CC HSSP; P02730; 1BTO.
CC InterPro; IPR001717; Anion_exchange.

Search completed: February 11, 2004, 17:04:16
Job time : 7.16667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 Seconds
(without alignments)
73.479 Million cell updates/sec

Title: US-09-901-187C-11
Perfect score: 44
Sequence: 1 THHPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_arChaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	95.5	468	10	081027
2	40	90.9	1202	5	Q81DT1
3	39	88.6	73	5	Q8Y32
4	39	88.6	312	10	Q9C7M4
5	39	88.6	402	2	Q53371
6	39	88.6	402	2	Q8VSK3
7	39	88.6	402	2	Q9K388
8	39	88.6	402	2	Q8RQK7
9	39	88.6	402	2	Q56368
10	39	88.6	402	2	Q8GC95
11	39	88.6	402	4	Q9H5N6
12	39	88.6	402	4	Q8BTP6
13	39	88.6	402	5	Q95TW0
14	39	88.6	402	5	Q95T96
15	39	88.6	402	12	Q8QTC0
16	39	88.6	402	12	Q9E8P4

17	39	88.6	402	16	Q98459
18	39	88.6	559	10	Q947U9
19	38	86.4	308	5	O76762
20	38	86.4	393	5	O8WSF5
21	38	86.4	393	5	O81537
22	38	86.4	470	10	O8LGX5
23	38	86.4	1377	5	Q9W2F3
24	37	84.1	66	16	Q9EDV2
25	37	84.1	69	16	Q9JYN1
26	37	84.1	112	4	O8N377
27	37	84.1	239	13	O8JHA2
28	37	84.1	260	2	Q93P79
29	37	84.1	343	13	Q9DG50
30	37	84.1	700	3	O8X007
31	37	84.1	723	5	Q9UB43
32	37	84.1	723	5	Q9W5F0
33	37	84.1	723	5	Q26436
34	37	84.1	840	11	Q8X062
35	36	81.8	156	16	Q9JUG8
36	36	81.8	184	16	O8VKL1
37	36	81.8	443	16	Q9JZC9
38	36	81.8	553	10	Q9ZK78
39	36	81.8	827	3	O60130
40	36	81.8	869	5	Q9Y0Z8
41	36	81.8	1041	5	O8WT17
42	36	81.8	1041	5	O9VHL6
43	35	79.5	169	16	Q9PEL8
44	35	79.5	250	16	Q912C2
45	35	79.5	271	16	Q8EFA4

ALIGNMENTS

RESULT 1

081027	081027	PRELIMINARY;	PRT;	468	AA.
AC	081027; Q9FPE2;				
DT	01-NOV-1998 (TRENBLrel. 08, Created)				
DT	01-JUN-2002 (TRENBLrel. 21, Last sequence update)				
DE	Putative hydroxymethylglutaryl-CoA lyase (Putative				
DE	hydroxymethylglutaryl-CoA lyase protein).				
GN	At226800.				
OS	Arabidopsis thaliana (Mouse-ear cress).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=cv. Columbia;				
RA	Rounsley S.D., Ronning C.M., Lin X., Ketchum K.A., Crosby M.L.,				
RA	Brandon R.C., Sykes S.M., Kaul S., Mason T.M., Kerlavage A.R.,				
RA	Adams M.D., Somerville C.R., Venter J.C.;				
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=cv. Columbia;				
RA	Lin X.;				
RA	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=cv. Columbia;				
RA	Town C.D., Kaul S.;				
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Sakano H., Liu S.X., Pham P.K., Yamada K., Banh J., Etgu P., Lee J.M.,				
RA	Toriumi M., Yu G., Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;				
RA	"Full Length cDNA of gene F12C20.16/At2g26800 (GI:3426048).";				
RT	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.				
RN	[5]				

SEQUENCE FROM N.A.

RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo C.J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RR EMBL; AY075431; AAL68246.1; -;
DR FLYBase; FBGN061414; BCNDA:IP02734.
SQ SEQUENCE 73 AA; 8338 MW; B1862F88AA421AA3 CRC64;

Query Match 88.6%; Score 39; DB 5; Length 73;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHHPS 7
 |||||
DB 39 HHHHPS 44

RESULT 4

Q9C7M4 PRELIMINARY; PRT; 312 AA.

ID AC Q9C7M4;
AC Q9C7M4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Bacterial IS-element.
DE F103.7.

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]

SEQUENCE FROM N.A.

RP STRAIN=cnv. Columbia;
RC MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltseher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana."
RL Nature 408:816-820(2000).
RR EMBL; AC068901; AAGS0897.1; -;
DR InterPro; IPR002559; Transposase.11.
DR Pfam; PF01609; Transposase.11; 1.
SQ SEQUENCE 312 AA; 35554 MW; 3C06A8F768CE29E8 CRC64;

Query Match 88.6%; Score 39; DB 10; Length 312;
Best Local Similarity 85.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THHHPS 7
 |||||
DB 157 THHHPS 163

```
RESULT 5
Q53371 PRELIMINARY; PRT; 402 AA.
AC Q53371;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Transposase (IS10 transposase).
GN R0085.
OS Escherichia coli, and
OS Salmonella typhi.
OG Plasmid pXT107.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562, 601;
RN [1]
RN SEQUENCE FROM N.A.
RP SPECIES=E.coli; PLASMID=pXT107;
RX MEDLINE=94040791; PubMed=8224890;
RA Bogosian G., Bilyeu K., O'Neill J.P.;
RT "Genome rearrangements by residual IS10 elements in strains of
RT Escherichia coli K-12 which had undergone Tn10 mutagenesis and fusaric
RT acid selection.";
RL Gene 133:17-22(1993).
RN [2]
RN SEQUENCE FROM N.A.
RP SPECIES=S.typhi; PLASMID=R27;
RX MEDLINE=20247426; PubMed=10773089;
RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,
RA Grobeck E., Rose D.J., Taylor D.E.;
RT "The complete DNA sequence and analysis of R27, a large IncHI plasmid
RT from Salmonella typhi that is temperature sensitive for transfer.";
RL Nucleic Acids Res. 28:2177-2186(2000).
RN [3]
RP SEQUENCE FROM N.A.
RP SPECIES=S.typhi; PLASMID=R27;
RX MEDLINE=20247426; PubMed=10783303;
RA Lawley T.D., Burland V.D., Taylor D.E.;
RT "Analysis of the complete nucleotide sequence of the tetracycline-
RT resistance transposon Tn10.";
RL Plasmid 43:235-239(2000).
DR EMBL; S67119; AAB28848.1; -.
DR EMBL; AF250878; AAF69923.1; -.
DR EMBL; AF223162; AAF69930.1; -.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Plasmid.
SQ SEQUENCE 402 AA; 46041 MW; 2B4A8708C885667C CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 247 THCHPS 253

RESULT 6
Q8VSK3 PRELIMINARY; PRT; 402 AA.
AC Q8VSK3;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Transposase.
GN CP0084.
OS Shigella flexneri 2a.
OG Plasmid pCP301.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OC NCBI_TaxID=42897;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=301;
RA Jin Q., Zhang J.Y., Liu H., Yang J., Yang F., Zhang X.B., Wang J.H.,
RA Yan G.W., Wu H.T., Dong J., Sun L.L., Xue Y., Zhao A.L., Gao Y.S.,
RA Zhu J.P., Kan B., Chen S.X., Yao Z.J., He B.K., Chen R.S., Ma D.L.,
RA Yuan Z.H., Xu J.G., Wang Y., Shen Y., Lu W.C., Qiang B.Q., Wen Y.M.,
Hou Y.D.;
RT "Complete DNA sequence and analysis of the large virulence plasmid
RT pCP301 of Shigella flexneri.";
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF386526; AAL72416.1; -.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Plasmid.
SQ SEQUENCE 402 AA; 46046 MW; 1E9391EE3244A12B6 CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 247 THCHPS 253

RESULT 7
Q9K388 PRELIMINARY; PRT; 402 AA.
AC Q9K388;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Putative IS10 transposase.
GN R0076.
OS Salmonella typhi.
OG Plasmid R27.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OC NCBI_TaxID=601;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=20280091; PubMed=10773089;
RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,
RA Grobeck E., Rose D.J., Taylor D.E.;
RT "The complete DNA sequence and analysis of R27, a large IncHI plasmid
RT from Salmonella typhi that is temperature sensitive for transfer.";
RL Nucleic Acids Res. 28:2177-2186(2000).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=20247426; PubMed=10783303;
RA Lawley T.D., Burland V.D., Taylor D.E.;
RT "Analysis of the complete nucleotide sequence of the tetracycline-
RT resistance transposon Tn10.";
RL Plasmid 43:235-239(2000).
DR EMBL; AF250878; AAF69914.1; -.
DR EMBL; AF223162; AAF69930.1; -.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Plasmid.
SQ SEQUENCE 402 AA; 46059 MW; 25854FABD793130F CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 247 THCHPS 253

RESULT 8
Q8RQK7 PRELIMINARY; PRT; 402 AA.
ID Q8RQK7
```


QROK7;
01-JUN-2002 (T-EMBLrel. 21, Created)
01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
Transposase.
GN TRPS.
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN=Tm-101;
RA Nishikawa M., Itoi Y., Ito S., Inoue M.;
RT "Genes expressed in Bacillus cereus.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB083420; BAB88926.1;
DR InterPro; IPR002559; Transposase 11.
DR Pfam; PF01609; Transposase 11; 1.
SQ SEQUENCE 402 AA; 46070 MW; 53D98F9BD09D5962 CRC64;
Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. NO. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 THHHPS 7
|||
Db 247 THCHPS 253
RESULT 9
Q56368 PRELIMINARY; PRT; 402 AA.
AC Q56368; Q9R2K0;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Hypothetical protein [Transposase of Tn10] (Tnpr protein).
GN RMP1 OR YEDA OR TNP.
OS Escherichia coli,
OS Methylobionas aminofaciens,
OS Shigella flexneri, and
OS Salmonella typhimurium.
OG Plasmid R64.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 46896, 623, 602;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; TRANSPOSON=Tn10;
RX MEDLINE=8222187; PubMed=6283536;
RA Halling S.M., Simons R.W., Way J.C., Walsh R.B., Kleckner N.;
RT "DNA sequence organization of IS10-right of Tn10 and comparison with
IS10-left.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2608-2612(1982).
RN [2]
RP PARTIAL SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; TRANSPOSON=Tn10;
RA Errade P.;
RL Submitted (APR-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.aminofaciens;
RX MEDLINE=99345693; PubMed=10418139;
RA Sakai Y., Mitsui R., Katayama Y., Yanase H., Kato N.;
RT "Organization of the genes involved in the ribulose monophosphate
pathway in an obligate methylotrophic bacterium, Methylobionas
aminofaciens 77a.";
RL FEMS Microbiol. Lett. 176:125-130(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; TRANSPOSON=Tn10;
RA Chalmers R., Crellin P., Sewitz S., Liplow K.;
RT "The complete nucleotide sequence of transposon Tn10.";

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; PLASMID=R100;
RA Sampei G., Mizobuchi K.;
RT "Organization and diversification of plasmid genomes: complete
nucleotide sequence of the R100 genome.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RA Sampei G., Komano T., Sasaki T., Tachibana K., Furuya N., Saito Y.,
RA Suzuki T., Mizobuchi K.;
RT "Organization and diversification of plasmid genomes: complete
nucleotide sequence of R64 genome.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=87146423; PubMed=3029698;
RA Komano T., Kubo A., Nishioka T.;
RT "Shuffle: multi-inversion of four contiguous DNA segments of plasmid
R64 creates seven different open reading frames.";
RL Nucleic Acids Res. 15:1165-1172(1987).
RN [8]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=89127142; PubMed=3065610;
RA Kubo A., Kusukawa A., Komano T.;
RT "Nucleotide sequence of the rci gene encoding shufflon-specific DNA
recombinase in the IncII plasmid R64: homology to the site-specific
Mol. Gen. Genet. 213:30-35(1988).
RN [9]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=88314948; PubMed=3045094;
RA Komano T., Toyoshima A., Morita K., Nishioka T.;
RT "Cloning and nucleotide sequence of the orit region of the IncII
plasmid R64.";
RL J. Bacteriol. 170:4385-4387(1988).
RN [10]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=91177811; PubMed=1848841;
RA Furuya N., Nishioka T., Komano T.;
RT "Nucleotide sequence and functions of the orit operon in IncII plasmid
R64.";
RL J. Bacteriol. 173:2231-2237(1991).
RN [11]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=92011438; PubMed=1917882;
RA Furuya N., Komano T.;
RT "Determination of the nick site at orit of IncII plasmid R64: global
similarity of orit structures of IncII and IncP plasmids.";
RL J. Bacteriol. 173:6612-6617(1991).
RN [12]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=93352408; PubMed=8349545;
RA Kim S.R., Runayama N., Komano T.;
RT "Nucleotide sequence and characterization of the traABCD region of
IncII plasmid R64.";
RL J. Bacteriol. 175:5035-5042(1993).
RN [13]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; PLASMID=R64;
RX MEDLINE=95083745; PubMed=7991676;
RA Furuya N., Komano T.;
RT "Surface exclusion gene of IncII plasmid R64: nucleotide sequence and
analysis of deletion mutants.";
RL Plasmid 32:80-84(1994).

[14]
SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
MEDLINE=96138149; PubMed=8626273;
Furuya N., Komano T.;
"Nucleotide sequence and characterization of ttrABC region of the
IncII plasmid R64: existence of the pnd gene for plasmid maintenance
within the transfer region.";
J. Bacteriol. 178:1491-1497(1996).
[15]
SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
MEDLINE=97428559; PubMed=9281491;
Narahara K., Rahman Z., Furuya N., Komano T.;
"Requirement of a limited segment of the sog gene for plasmid R64
cojugation.";
Plasmid 38:1-11(1997).
[16]
SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
MEDLINE=97315231; PubMed=9171405;
Kim S.R., Komano T.;
"The plasmid R64 thin pilus identified as a type IV pilus.";
J. Bacteriol. 179:3594-3603(1997).
[17]
SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
MEDLINE=98053841; PubMed=9393692;
Furuya N., Komano T.;
"Mutational analysis of the R64 oriT region: requirement for precise
location of the NKA-binding sequence.";
J. Bacteriol. 179:7291-7297(1997).
[18]
SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
MEDLINE=98268996; PubMed=9603870;
Yoshida T., Furuya N., Ishikura M., Isobe T., Haino-Fukushima K.,
Ogawa T., Komano T.;
"Purification and characterization of thin pili of IncII plasmids
Colib-P9 and R64: formation of PiliV-specific cell aggregates by type
IV pili.";
J. Bacteriol. 180:2842-2848(1998).
[19]
SEQUENCE FROM N.A.
SPECIES=S.typhimurium; PLASMID=R64;
MEDLINE=20223621; PubMed=10760136;
Komano T., Yoshida T., Narahara K., Furuya N.;
"The transfer region of IncII plasmid R64: similarities between R64
tra genes and Legionella icm/dot genes.";
Mol. Microbiol. 35:1348-1359(2000).
DR EMBL; J01829; AAA88660.1; -;
DR EMBL; AB026428; BAA83097.1; -;
DR EMBL; AF162223; AAD50250.1; -;
DR EMBL; AP000342; BAA78838.1; -;
DR EMBL; AP005147; BAB91573.1; -;
DR InterPro; IPR002559; Transposase_11;
DR Pfam; PF01609; Transposase_11; 1;
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 402 AA; 46027 MW; 25854FABC8857EDF CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 247 THCHPS 253

RESULT 10
Q8GC95 PRELIMINARY; PRT; 402 AA.
AC Q8GC95;

OS Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

DR 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Transposase.
GN TNP.
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=546;
RN [1] |||||
RP SEQUENCE FROM N.A.
RC STRAIN=3009;
RA Oelschlaeger T.A.;
RT "Cloning, sequencing and expression of an invasion determinant from
Citrobacter freundii strain 3009.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ508060; CAD56977.1; -;
SQ SEQUENCE 402 AA; 46027 MW; 25854FABC8857EDF CRC64;

Query Match 88.6%; Score 39; DB 2; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 247 THCHPS 253

RESULT 11
Q9HSN6 PRELIMINARY; PRT; 402 AA.
ID Q9HSN6;
AC Q9HSN6;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Hypothetical protein FLJ23250.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] |||||
RP SEQUENCE FROM N.A.
RC Tissue=Colon;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
Tanaka T., Nakamura Y., Isogai T., Sugano S.;
"NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK026903; BAB15586.1; -;
DR InterPro; IPR002559; Transposase_11;
DR Pfam; PF01609; Transposase_11; 1;
KW Hypothetical protein.
SQ SEQUENCE 402 AA; 46027 MW; 25854FABC8857EDF CRC64;

Query Match 88.6%; Score 39; DB 4; Length 402;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7
Db 247 THCHPS 253

RESULT 12
Q9BTF6 PRELIMINARY; PRT; 402 AA.
ID Q9BTF6;
AC Q9BTF6;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Search completed: February 11, 2004, 17:09:46
Job time : 26.5833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 Seconds
(without alignments)
73.479 Million cell updates/sec

Title: US-09-901-187C-12
Perfect score: 34
Sequence: 1 M M M M M R L 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phages:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	94.1	70	Q9LZD2	Q9LZD2 arabidopsis
2	30	88.2	192	Q9SQN9	Q9SQN9 arabidopsis
3	30	88.2	249	Q8LDG1	Q8LDG1 arabidopsis
4	30	88.2	249	Q9SVL0	Q9SVL0 arabidopsis
5	30	88.2	353	Q9ST57	Q9ST57 arabidopsis
6	30	88.2	364	Q9AWZ8	Q9AWZ8 oryza sativ
7	28	82.4	129	Q9SRT0	Q9SRT0 drosophila
8	28	82.4	307	Q8WA93	Q8WA93 narceus ann
9	28	82.4	307	Q8ZM25	Q8ZM25 salmonella
10	27	79.4	39	O00868	O00868 plasmodium
11	27	79.4	70	Q8L823	Q8L823 arabidopsis
12	27	79.4	103	Q8H4V8	Q8H4V8 oryza sativ
13	27	79.4	147	Q9S870	Q9S870 plasmodium
14	27	79.4	176	Q8WBQ4	Q8WBQ4 apis labori
15	27	79.4	238	Q9FKE0	Q9FKE0 arabidopsis
16	27	79.4	283	Q9IQD9	Q9IQD9 drosophila

Q8RUX0 oryza sativ
Q8M355 saccharomyc
Q9KH45 pantoaea agg
Q9S712 arabidopsis
Q9NBX5 drosophila
Q8LHX4 oryza sativ
Q8LXC2 plasmodium
Q8R2J0 mus musculu
Q8LJF5 plasmodium
Q9H3F1 homo sapien
Q8N6W1 homo sapien
Q8PB18 xanthomonas
Q8PB41 xanthomonas
Q9PDG3 xylella fas
Q9LJY9 arabidopsis
Q8H10 melitaea pe
Q9M8E6 arabidopsis
Q9K953 bacillus ha
Q8VGM0 mus musculu
Q9S7Q0 arabidopsis
Q9NG0 bombyx mori
Q9SR3 arabidopsis
Q9N55 anopheles g
Q8WSX8 anopheles g
Q9FCY8 erwinia ste
Q9C954 arabidopsis
Q45007 caenorhabdi
Q92VF2 rhizobium m
Q9GCJ9 parides pho

ALIGNMENTS

RESULT 1

Q9LZD2 Q9LZD2 PRELIMINARY; PRT; 70 AA.
ID Q9LZD2
AC Q9LZD2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 8.0 kDa protein.
GN F12E4330.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Dale J.M., Gibson H.A., Goldsmith A.D., Jiang P.X., Lee J.M.,
RA Quach H.L., Tang C., Toriumi M., Yu G., Bowser L., Carninci P.,
RA Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Kim C., Kosemura E., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shimizu P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RL "Full Length cDNA of gene F12E4.330 (GI:7378640).";
RN Submitted (JUN-2001) to the EMBL/GenBank/DBSJ databases.
RL EMBL; AL162751; CAB83316.1; -
DR EMBL; AT040018; AAK6175.1; -
KW Hypothetical protein.
SQ SEQUENCE 70 AA; 8024 MW; A224B96F68370BC7 CRC64;

```

Query Match      94.1%; Score 32; DB 10; Length 70;
Best Local Similarity 85.7%; Pred. No. 1.4;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWMWML 7
Db 3 MWMWML 9

RESULT 2
Q9SQN9 PRELIMINARY; PRT; 192 AA.
AC Q9SQN9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 21.2 kDa protein.
GN F19K16.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Uterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RA "Arabidopsis thaliana chromosome 1 BAC F19K16 genomic sequence.";
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A.
RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,
RA Shinn P., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RA "Genomic sequence for Arabidopsis thaliana BAC F20B17 from chromosome
RT I.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011717; AAC52257.1; -.
DR EMBL; AC010793; AAF68112.1; -.
DR HSSP; P29602; 1JER.
DR InterPro; IPR003245; Plicyanin_like.
DR Pfam; PF02298; Cu_bind_like; 1.
DR ProDom; PD003122; Plicyanin_like; 1.
KW Hypothetical protein.
SQ SEQUENCE 192 AA; 21197 MW; 2322FD669A95A04C CRC64;

Query Match      88.2%; Score 30; DB 10; Length 192;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWMWML 6
Db 2 MWMWML 7

RESULT 3
Q9SVL0 PRELIMINARY; PRT; 249 AA.
AC Q9SVL0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 28.7 kDa protein.
GN F18B3.170.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Rieger M., Gabel C., Mueller-Auer S., Schaefer M., Zipp M.,
RA Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C., Quetier F.,
RA Salanoubat M.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL049862; CAB42918.1; -.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR006455; Homeo_ZF_HD.
DR InterPro; IPR006456; ZF_HD_N.
DR InterPro; IPR006857; ZNF_HD_dimer.

```

DR Pfam; PF04770; ZF-HD dimer; 1.
 DR TIGRFAMs; TIGR01565; homeo_ZF_HD; 1.
 DR TIGRFAMs; TIGR01566; ZF_HD_prot_N; 1.
 DR PROSITE; PS00071; HOMEBOX_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 249 AA; 28658 MW; F0C163A282D51BB4 CRC64;
 Query Match 88.2%; Score 30; DB 10; Length 249;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MWMWMMR 6
 Db 174 MWMWMMR 179

RESULT 5
 Q9ST57 ID Q9ST57 PRELIMINARY; PRT; 353 AA.
 AC Q9ST57
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 39.1 kDa protein.
 GN T23J7.190.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsi.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Choise N., Robert C., Bröttier P., Wincker P., Cattolico L.,
 RA Attiguenave F., Saurin C., Weissenbach J., Mewes H.W., Mayer K.F.X.,
 RA Lemcke K., Schueller C., Queirer F., Salancubet M.,
 RA Yamada K., Ecker J., Theologis A., Davis R.W.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Chen H., Cheuk R.,
 RA Chung M.K., Kim C., Lin J., Liu S.X., Pham P.K., Sakano H., Shinn P.,
 RA Yamada K., Ecker J., Theologis A., Davis R.W.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
 RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RL "Full Length cDNA of Gene T23J7.190/AT947860 (GI:4741203).";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL049746; CAB41869.1; -
 DR EMBL; AY044334; AAK73275.1; -
 DR EMBL; AY035165; AAK59669.1; -
 DR InterPro; IPR001064; Crystallin.
 DR InterPro; IPR000566; Lipocalin_cytFABP.
 DR Pfam; PF00061; lipocalin; 1.
 DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
 DR PROSITE; PS00213; LIPOCALIN; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 353 AA; 39116 MW; 14A83B0CF1BA8D0D CRC64;
 Query Match 88.2%; Score 30; DB 10; Length 353;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWMWMMR 6
 Db 122 MWMWMMR 127

RESULT 6
 Q9AWZ8 ID Q9AWZ8 PRELIMINARY; PRT; 364 AA.
 AC Q9AWZ8
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE 3-methyl-2-oxobutanoate hydroxy-methyl-transferase-like protein.
 GN P0492F05.14 OR P0443E07.6.
 OS Oryza sativa (Rice), and
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OC NCBI_TaxID=4530, 39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPSCIES=O.sativa; STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 clone: P0492F05.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPSCIES=O.sativa (japonica cultivar-group); STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 clone: P0443E07.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP002902; BAB32712.1; -
 DR EMBL; AP002900; BAB92103.1; -
 DR Gramene; Q9AWZ8;
 DR InterPro; IPR003700; Pantoate transf.
 DR Pfam; PF02548; Pantoate_transf; 1.
 DR TIGRFAMs; TIGR00222; panB; 1.
 KW Transferase.
 SQ SEQUENCE 364 AA; 37920 MW; 3393COA29F2B1512 CRC64;
 Query Match 88.2%; Score 30; DB 10; Length 364;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWMWMMR 6
 Db 1 MWMWMMR 6

RESULT 7
 Q95RTO ID Q95RTO PRELIMINARY; PRT; 129 AA.
 AC Q95RTO
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE LD12764P.
 GN BCDNA:LD12764.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunco J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,

```

RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY061159; AAL28707.1; -;
DR Flybase; Fgsc0047212; BCNA.LD.2764.
SQ SEQUENCE 129 AA; 14240 MW; C349AABE9EC3DBEF CRC64;

Query Match      82.4%; Score 28; DB 5; Length 129;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNNMMRL 7
DB 13 MNNMMRL 19

RESULT 8
Q8WA93 PRELIMINARY; PRT; 307 AA.
ID Q8WA93
AC Q8WA93; 2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE NADH dehydrogenase subunit 1 (EC 1.6.5.3) (NADH-ubiquinone
DE oxidoreductase chain 1).
GN NAD1
OS Nardus annularis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Diplopoda; Helminthomorpha;
OC Spirobolidae; Nardus.
OX NCBI_TaxID=174156;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21660208; PubMed=11801744;
RA Lavrov D.V., Boore J.L., Brown W.M.;
RT "Complete mtDNA Sequences of Two Millipedes Suggest a New Model for
RT Mitochondrial Gene Rearrangements: Duplication and Nonrandom Loss.";
RL Mol. Biol. Evol. 19:163-169(2002).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=21660208; PubMed=11801744;
RA Lavrov D.V.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
DR EMBL; AY055727; AAL18214.1; -;
DR InterPro; IPR001694; Resp_NADH_dh1.
DR Pfam; PF00146; NADHdh; 1.
DR PROSITE; PS00667; COMPLEX1_ND1_1; 1.
DR PROSITE; PS00668; COMPLEX1_ND1_2; 1.
KW NAD; Oxidoreductase; Transmembrane; Ubiquinone; Mitochondrion.
SQ SEQUENCE 307 AA; 34669 MW; 03F837A142D8A24D CRC64;

Query Match      82.4%; Score 28; DB 8; Length 307;
Best Local Similarity 85.7%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNNMMRL 7
DB 233 MNNMMRL 239

RESULT 9
Q8ZM25 PRELIMINARY; PRT; 307 AA.
ID Q8ZM25
AC Q8ZM25; 2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative xylanase/chitin deacetylase.
GN STM3132
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;

[1]
RN SEQUENCE FROM N.A.
RX STRAIN=L72 / SGSC1412 / ATCC 700720;
RC MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florescu L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
DR EMBL; AE008844; AAL22006.1; -;
DR InterPro; IPR002509; Polysac_deact.
DR Pfam; PF01522; Polysac_deact; 1.
KW Xylan degradation; Hydrolase; Glycosidase; Hypothetical protein;
KW Complete proteome.
SQ SEQUENCE 307 AA; 35194 MW; B14436094A935245 CRC64;

Query Match      82.4%; Score 28; DB 16; Length 307;
Best Local Similarity 71.4%; Pred. No. 47;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNMMRL 7
DB 249 MNNMMRL 255

RESULT 10
Q00868 PRELIMINARY; PRT; 39 AA.
ID Q00868
AC Q00868;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE HMG-like protein Pf16 (Fragment).
GN PF16.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
[1]
RN SEQUENCE FROM N.A.
RX STRAIN=HB3;
RC MEDLINE=92118046; PubMed=1731798;
RA Guntaka R.V., Kandala J.C., Reddy V.D.;
RT "Cloning and characterization of a highly conserved HMG-like protein
RT (PF16) gene from Plasmodium falciparum.";
RL Biochem. Biophys. Res. Commun. 182:412-419(1992).
[2]
RN SEQUENCE FROM N.A.
RX STRAIN=HB3;
RC Nambiar A., Kandala J.C., Dolan S.A., Jensen J.J., Guntaka R.V.;
RT "Molecular cloning and characterization of a cDNA for the highly
RT conserved HMG-like protein (Pf16) gene of Plasmodium falciparum.";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U97128; AAB58116.1; -;
FT NON TER 39
SQ SEQUENCE 39 AA; 5037 MW; 012D80DD1F566823 CRC64;

Query Match      79.4%; Score 27; DB 5; Length 39;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNMMRL 6
DB 21 MNNMMRL 26

RESULT 11
Q8LE23 PRELIMINARY; PRT; 70 AA.
ID Q8LE23
AC Q8LE23;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)

```


DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
GN Hypothetical protein.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eubasids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volkovskiy N., Town C.D., Troughan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.,
RA "Full-length messenger RNA sequences greatly improve genome
RT annotation.", 0:0-0(2002).
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troughan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.,
RA "Full-length cDNA from Arabidopsis thaliana.",
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085662; AAW67306.1; -
KW Hypothetical protein.
SQ SEQUENCE 70 AA; 7270 MW; 10C0764E0986E031 CRC64;
Query Match 79.4%; Score 27; DB 10; Length 70;
Best Local Similarity 83.3%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWMWMM 6
DB 1 MWMWMM 6
RESULT 12
ID Q8H4V8 PRELIMINARY; PRT; 103 AA.
AC Q8H4V8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE OJ1484_G09.19 protein.
GN OJ1484_G09.19,
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriactoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
RT clone:OJ1484_G09.",
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003913; BAC24906.1; -
SQ SEQUENCE 103 AA; 11753 MW; 22AB780BF89FCB0 CRC64;
Query Match 79.4%; Score 27; DB 10; Length 103;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWMWMM 7
DB 45 MWMWMM 51
RESULT 13
ID Q25870 PRELIMINARY; PRT; 147 AA.
AC Q25870;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE High mobility group-like protein.
GN PS16.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92118046; PubMed=1731798;
RA Clontak R.V., Kandaia J.C., Vudem D.,
RA "Cloning and characterization of a highly conserved HMG-like protein
RT (PF16) gene from Plasmodium falciparum.",
RL Biochem. Biophys. Res. Commun. 182:412-419(1992).
DR EMBL; M86518; AAA29615.1; -
SQ SEQUENCE 147 AA; 17175 MW; A46C5D47A62DF75A CRC64;
Query Match 79.4%; Score 27; DB 5; Length 147;
Best Local Similarity 83.3%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWMWMM 6
DB 11 MWMWMM 16
RESULT 14
ID Q8WBQ4 PRELIMINARY; PRT; 176 AA.
AC Q8WBQ4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE NADH dehydrogenase subunit 4 (fragment).
GN ND4.
OS Apis laboriosa (Himalayan honeybee).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae; Apis.
OX NCBI_TaxID=183418;
RN [1]
RP SEQUENCE FROM N.A.
RA Leelamanit W., Amano K.;
RA "The NADH dehydrogenases of Apis mellifera, A. cerana, A. dorsata, A.
RT laboriosa and A. florea: Sequence Comparison and Genetic Diversity.",
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF47064; AAL60208.1; -
DR InterPro; IPR003918; NADHoxred4.
DR InterPro; IPR001750; Oxidored4.
DR Pfam; PF00361; Oxidored4.
DR PRINTS; PR01437; NUOXDRDTASE4.
DR NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
KW NON_TER 1
FT NON_TER 176
SQ SEQUENCE 176 AA; 19981 MW; 9E405EED9CF06687 CRC64;
Query Match 79.4%; Score 27; DB 8; Length 176;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWMWMM 6
DB 34 MWMWMM 39
RESULT 15
ID Q9FKE0 PRELIMINARY; PRT; 238 AA.
AC Q9FKE0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GB|AAF30306.1.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Ericaceae; II: Brassicales; Brassicaceae; Arabidopsis.
 CX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=98403884; PubMed=9734815;
 RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
 RT Sequence features of the regions of 1,367,185 bp covered by 19
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:203-216(1998).
 DR ENEL; AB012241; BAB09032.1; -;
 DR InterPro; IPR006121; HeavyMe transpt.
 DR InterPro; IPR006191; Metal_bind.
 DR Pfam; PF00403; HMA_1
 DR PROSITE; PS00846; HMA_2; 1.
 SQ SEQUENCE 238 AA; 27717 MW; F2F486C355011632 CRC64;

Query Match 79.4%; Score 27; DB 10; Length 238;
 Best Local Similarity 71.4%; Pred. No. 66;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAMMRL 7
 Db 103 MAMMRL 109

Search completed: February 11, 2004, 17:09:48
 Job time : 26.5833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
(without alignments)
63.714 Million cell updates/sec

Title: US-09-901-187C-12
Perfect score: 34
Sequence: 1 MMMXRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	29	85.3	132	1 F802 SCHWA	P16464 schistosoma
2	27	79.4	146	1 2SS_BEREX	P04403 bertholletii
3	27	79.4	1610	1 CCAD MEGAU	Q99244 mesocricetu
4	27	79.4	2161	1 CCAD HUMAN	Q01668 homo sapien
5	27	79.4	2203	1 CCAD RAT	R27732 rattus norv
6	26	76.5	395	1 RIF2 YEAST	Q06208 saccharomyc
7	26	76.5	403	1 HRPN_ERWAM	Q01039 erwania amy
8	26	76.5	438	1 POP1 CAEEL	Q10666 caenorhabdi
9	26	76.5	875	1 AMD2 XENLA	P12890 xenopus lae
10	25	73.5	313	1 NU1M_RHISA	Q99824 rhinicephal
11	25	73.5	321	1 AOX1_SOVEN	Q07185 glycine max
12	25	73.5	325	1 CV1_HUMAN	P08574 homo sapien
13	25	73.5	355	1 HXL6 LYCES	Q22299 lycopersico
14	25	73.5	377	1 YAWD_RHISN	P55682 rhizobium s
15	25	73.5	382	1 STM_ARATH	Q38874 arabidopsis
16	25	73.5	383	1 STM_BRAOL	Q96609 brassica ol
17	25	73.5	400	1 AMD1 XENLA	P08478 xenopus lae
18	25	73.5	457	1 NORM_ECO57	P58164 escherichia
19	25	73.5	457	1 NORM_ECOLI	P37340 escherichia
20	25	73.5	554	1 NUSM_APILL	P34855 apis mellif
21	25	73.5	783	1 YKR2 CAEEL	P34308 caenorhabdi
22	25	73.5	931	1 DING_BACSU	P54394 bacillus su
23	24	70.6	96	1 CLV3_ARATH	Q9x104 arabidopsis
24	24	70.6	99	1 PD11_MOUSE	Q9y691 homo sapien
25	24	70.6	99	1 P11_MOUSE	P56983 mus musculu
26	24	70.6	179	1 Y281_METJA	Q57729 methanococc
27	24	70.6	181	1 G12_PSEAE	Q9i351 pseudomonas
28	24	70.6	186	1 G11_PSEAE	Q9hy98 pseudomonas
29	24	70.6	265	1 CTRC_NEIMA	P57012 neisseria m
30	24	70.6	322	1 PYRB_XYLFA	Q9pbb8 xyliella fas
31	24	70.6	423	1 MB12 YEAST	P03873 saccharomyc
32	24	70.6	434	1 AVT_CATCO	Q90352 catostomus
33	24	70.6	437	1 PUR2_XYLFA	Q9pc09 xyliella fas

34	24	70.6	462	1 ZRAS_KLEOX	Q9ape0 klebsiella
35	24	70.6	503	1 CP51_HUMAN	Q16850 homo sapien
36	24	70.6	562	1 ILVD_PYRAE	Q8zyu6 pyrobaculum
37	24	70.6	582	1 YC73_MYCTU	Q11046 mycobacteri
38	24	70.6	687	1 SRB4_YEAST	P32569 saccharomyc
39	24	70.6	902	1 MOPT_ARATH	P93024 arabidopsis
40	24	70.6	1579	1 SSK2 YEAST	P53599 saccharomyc
41	23	67.6	49	1 Y495_TREPA	O83508 treponema p
42	23	67.6	72	1 YNEF_BACSU	P45708 bacillus su
43	23	67.6	75	1 YA26_STAEP	O8csp3 staphylococ
44	23	67.6	80	1 YD43_STAAM	O99ud3 staphylococ
45	23	67.6	89	1 Y703_ARCFU	O29555 archaeoglob

ALIGNMENTS

RESULT 1

F802 SCHWA

ID F802 SCHWA STANDARD; PRT; 132 AA.

AC P16464;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Female specific 800 protein (FS800).

OS Schistosoma mansoni (Blood fluke).

OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;

OC Schistosomatidae; Schistosomatidae; Schistosoma.

OX NCBI_TaxID=6183;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Puerto Rican;

RX MEDLINE=89181810; PubMed=2927441;

RT Reiss M.G., Kuhn J., Blanton R., Davis A.H.;

RT "Localization and pattern of expression of a female specific mRNA in

RT Schistosoma mansoni.";

RL Mol. Biochem. Parasitol. 32:113-119(1989).

CC -!- FUNCTION: FS800 is likely to have some function in the production

CC or maintenance of the schistosome egg. It may have a function

CC unrelated to eggshell formation.

CC -!- DEVELOPMENTAL STAGE: Highest level only in mature worms, i.e.,

CC during egg production.

CC -!- MISCELLANEOUS: The two F800 proteins are read from two

CC overlapping reading frames.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL, J03999; AAA29884.1; -.

FT DOMAIN 2 21 ILE-RICH.

FT DOMAIN 47 53 POLY-MET.

SQ SEQUENCE 132 AA; 15561 MW; 02C77F42A25E120E CRC64;

Query Match 85.3%; Score 29; DB 1; Length 132;

Best Local Similarity 71.4%; Pred. No. 2;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMXRL 7

Db 49 MMMXMKI 55

|||||:

RESULT 2

2SS_BEREX

ID 2SS_BEREX STANDARD; PRT; 146 AA.

AC P04403; P04402;

DT 20-MAR-1987 (Rel. 04, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 2S sulfur-rich seed storage protein precursor (Allergen Ber e 1).
GN B2S1 AND B2S2.
OS Bertholletia excelsa (Brazil nut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Aceridae; Ericales; Lecythidaceae; Bertholletia.
OC NCBI_TaxID=3645;
RN [1]
RN SEQUENCE FROM N.A.
RP Altenbach S.B., Pearson K.W., Leung F.W., Sun S.S.M.;
RA "Cloning and sequence analysis of a cDNA encoding a Brazil nut protein
RT exceptionally rich in methionine";
RL Plant Mol. Biol. 8:239-250(1987).
RN [2]
RN SEQUENCE FROM N.A.
RP Bassuener R.;
RA Submitted (DEC-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RP MEDLINE=91370890; PubMed=1840683;
RX Gander E.S., Holmstrom K.O., de Paiva G.R., de Castro L.A.B.,
RA Carneiro M., Grossi de Sa M.F.;
RA "Isolation, characterization and expression of a gene coding for a 2S
RT albumin from Bertholletia excelsa (Brazil nut).";
RL Plant Mol. Biol. 16:437-448(1991).
RN [4]
RN SEQUENCE OF 37-64 AND 70-142
RX MEDLINE=87004679; PubMed=3758080;
RA Ampe C., van Damme J., de Castro L.A.B., Sampaio M.J.A.M.,
RA van Montagu M., Vandeckerckhove J.;
RT "The amino-acid sequence of the 2S sulphur-rich proteins from seeds
RT of Brazil nut (Bertholletia excelsa H.B.K.).";
RL Eur. J. Biochem. 159:597-604(1986).
CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
CC LINKED BY DISULFIDE BONDS.
CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M17146; AAA33010.1; -
DR EMBL; X57027; CAA40343.1; -
DR EMBL; X57028; CAA40344.1; -
DR EMBL; X54490; CAA38362.1; -
DR EMBL; X54491; CAA38363.1; ALT_SEQ.
DR EMBL; A13818; CAA01131.1; -
DR PIR; A25802; A25802.
DR PIR; S14946; S14946.
DR PDB; 1GYS; 30-APR-02.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR000817; Nepin.
DR Pfam; PF00234; tryp_alpha_aml; 1.
DR PRINTS; PRO0496; NAFIN.
DR SMART; SM00499; AAI; 1.
DR Seed storage protein; Signal; Allergen; Pyrrolidone carboxylic acid;
KW 3D-structure. 1 22
FT SIGNAL 23 36
FT PROPEP 37 64 SMALL CHAIN.
FT CHAIN 65 69
FT PROPEP 70 142 LARGE CHAIN 1B.
FT CHAIN 143 146
FT PROPEP 37 37 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 91 91 S -> E (IN VARIANT 1A).
FT VARIANT 28 29 EE -> QQ (IN REF. 4).
FT CONFLICT 122 122 L -> M (IN REF. 4).
FT

FT CONFLICT 126 126 I -> L (IN REF. 4).
SQ SEQUENCE 146 AA; 16911 MW; A7DF778FD766410D CRC64;
Query Match 79.4%; Score 27; DB 1; Length 146;
Best Local Similarity 83.3%; Pred. No. 6.7;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 MWMWRL 7
DB 99 MWMWRL 104
RESULT 3
CCAD MESAU
ID_CCAD MESAU STANDARD; PRT; 1610 AA.
AC Q99244; Q99245;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Voltage-dependent L-type calcium channel alpha-1D subunit (Calcium
DE channel, L type, alpha-1 polypeptide isoform 2).
GN CACNA1D OR CACNA1A2 OR CCHL1A2 OR CACH3 OR CACNA4.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM HCA3A).
RX TISSUE=Insulinoma;
RX MEDLINE=93149124; PubMed=1337146;
RA Yancy G.C., Wheeler M.B., Wei X., Perez-Reyes E., Birnbaumer L.,
RA Boyd A.E. III, Moss L.G.;
RA "Cloning of a novel alpha 1-subunit of the voltage-dependent calcium
RT channel from the beta-cell.";
RL Mol. Endocrinol. 6:2143-2152(1992).
RN [2]
RN SEQUENCE OF 1146-1441 FROM N.A. (ISOFORMS CACH3B AND CACH3D).
RP TISSUE=Heart;
RX MEDLINE=91056091; PubMed=2173707;
RA Perez-Reyes E., Wei X., Castellano A., Birnbaumer L.;
RT "Molecular diversity of L-type calcium channels. Evidence for
RT alternative splicing of the transcripts of three non-allelic genes.";
RL J. Biol. Chem. 265:20430-20436(1990).
CC -1- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1D
CC GIVES RISE TO L-TYPE CALCIUM CURRENTS. LONG-LASTING (L-TYPE)
CC CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)
CC GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP),
CC PHENYLALKYLAMINES, BENZOTHAZEPINES, AND BY OMEGA-AGATOXIN-IIIa
CC (OMEGA-AGA-IIIa). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-
CC GVIA (OMEGA-CTX-GVIA) AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA).
CC -1- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=HCA3A;
CC IsoId=Q99244-1; Sequence=Displayed;
CC Name=CACH3B;
CC IsoId=Q99244-3; Sequence=Not described;
CC Name=CACH3D;
CC IsoId=Q99244-2; Sequence=VSP 000915;
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, HEART AND SKELETAL MUSCLE.

RP SEQUENCE FROM N.A. (ISOFORM BETA-CELL-TYPE).
 RC TISSUE=Pancratic islets;
 RX MEDLINE=92115705; PubMed=1309948;
 RA Seino S., Chen L., Seino M., Blondel O., Takeda J., Johnson J.H.,
 RA Bell G.;
 RT "Cloning of the alpha 1 subunit of a voltage-dependent calcium
 RT channel expressed in pancreatic beta cells";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:584-588(1992).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM BETA-CELL-TYPE), AND VARIANT NIDDM.
 RX MEDLINE=9604438; PubMed=7557998;
 RA Yamada Y., Masuda K., Li Q., Ihara Y., Kubota A., Miura T.,
 RA Nakamura K., Fujii Y., Seino S., Seino Y.;
 RT "The structures of the human calcium channel alpha 1 subunit
 RT (CACNL1A2) and beta subunit (CACNLB3) genes";
 RL Genomics 27:312-319(1995).
 CC -!- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
 CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
 CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
 CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
 CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1D
 CC CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)
 CC GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP),
 CC PHENALKYLAMINES, BENZOTHAZEPINES, AND BY OMEGA-AGATOXIN-IIIA
 CC (OMEGA-AGA-IIIA). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-
 CC GVIA (OMEGA-CTG-GVIA) AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA).
 CC -!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
 CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
 CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
 CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
 CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
 CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
 CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Comment-Additional isoforms seem to exist;
 CC Name=Neuronal-type;
 CC IsoId=001668-1; Sequence=Displayed;
 CC Name=Beta-cell-type;
 CC IsoId=001668-2;
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN PANCREATIC ISLETS AND IN BRAIN,
 CC WHERE IT HAS BEEN SEEN IN HIPPOCAMPUS, BASAL GANGLIA, HYPOTHALAMUS AND
 CC THALAMUS. NO EXPRESSION IN SKELETAL MUSCLE.
 CC -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
 CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
 CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
 CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
 CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
 CC -!- POLYMORPHISM: A CHANGE FROM SEVEN TO EIGHT ATG TRINUCLEOTIDE
 CC REPEATS, RESULTING IN AN ADDITIONAL N-TERMINAL METHIONINE, HAS
 CC BEEN FOUND IN A PATIENT WITH NON-INSULIN-DEPENDENT DIABETES
 CC MELLITUS (NIDDM).
 CC -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M76558; AAA58402.1; -
 CC EMBL; M83566; AAA35629.1; -
 CC EMBL; D43747; BAA07804.1; -
 CC EMBL; D43705; BAA07804.1; JOINED.
 CC EMBL; D43706; BAA07804.1; JOINED.
 CC EMBL; D43707; BAA07804.1; JOINED.
 CC EMBL; D43708; BAA07804.1; JOINED.
 CC EMBL; D43709; BAA07804.1; JOINED.

DR EMBL; D43710; BAA07804.1; JOINED.
 DR EMBL; D43711; BAA07804.1; JOINED.
 DR EMBL; D43712; BAA07804.1; JOINED.
 DR EMBL; D43713; BAA07804.1; JOINED.
 DR EMBL; D43714; BAA07804.1; JOINED.
 DR EMBL; D43715; BAA07804.1; JOINED.
 DR EMBL; D43716; BAA07804.1; JOINED.
 DR EMBL; D43717; BAA07804.1; JOINED.
 DR EMBL; D43718; BAA07804.1; JOINED.
 DR EMBL; D43719; BAA07804.1; JOINED.
 DR EMBL; D43720; BAA07804.1; JOINED.
 DR EMBL; D43721; BAA07804.1; JOINED.
 DR EMBL; D43722; BAA07804.1; JOINED.
 DR EMBL; D43723; BAA07804.1; JOINED.
 DR EMBL; D43724; BAA07804.1; JOINED.
 DR EMBL; D43725; BAA07804.1; JOINED.
 DR EMBL; D43726; BAA07804.1; JOINED.
 DR EMBL; D43727; BAA07804.1; JOINED.
 DR EMBL; D43728; BAA07804.1; JOINED.
 DR EMBL; D43729; BAA07804.1; JOINED.
 DR EMBL; D43730; BAA07804.1; JOINED.
 DR EMBL; D43731; BAA07804.1; JOINED.
 DR EMBL; D43732; BAA07804.1; JOINED.
 DR EMBL; D43733; BAA07804.1; JOINED.
 DR EMBL; D43734; BAA07804.1; JOINED.
 DR EMBL; D43735; BAA07804.1; JOINED.
 DR EMBL; D43736; BAA07804.1; JOINED.
 DR EMBL; D43737; BAA07804.1; JOINED.
 DR EMBL; D43738; BAA07804.1; JOINED.
 DR EMBL; D43739; BAA07804.1; JOINED.
 DR EMBL; D43740; BAA07804.1; JOINED.
 DR EMBL; D43741; BAA07804.1; JOINED.
 DR EMBL; D43742; BAA07804.1; JOINED.
 DR EMBL; D43743; BAA07804.1; JOINED.
 DR EMBL; D43744; BAA07804.1; JOINED.
 DR EMBL; D43745; BAA07804.1; JOINED.
 DR EMBL; D43746; BAA07804.1; JOINED.
 DR Genew; HGNC:1391; CACNALD.
 DR NIM; 114206; -
 DR GO; GO:0005891; C:voltage-gated calcium channel complex; TAS.
 DR GO; GO:0015270; F:dihydropyridine-sensitive calcium channel a. . .; TAS.
 DR GO; GO:0006832; P:small molecule transport; TAS.
 DR InterPro; IPR001682; Ca/Na_pore.
 DR InterPro; IPR002077; Ca channel.
 DR InterPro; IPR002111; Cat channel_TripL.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR005446; LVDCALPHAL.
 DR InterPro; IPR005452; LVDCALPHALD.
 DR InterPro; IPR005820; M+channel_nig.
 DR Pfam; PF00520; ion_trans_4.
 DR PRINTS; PRO167; CACHANNEL.
 DR PRINTS; PRO1630; LVDCALPHAL.
 DR PRINTS; PRO1636; LVDCALPHALD.
 DR Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
 KW Calcium channel; Glycoprotein; Repeat; Multigene family;
 KW Calcium-binding; Phosphorylation; Alternative splicing; Polymorphism;
 KW Triplet repeat expansion.
 FT REPEAT 113 409
 FT REPEAT 113 409
 FT REPEAT 509 755
 FT REPEAT 873 1155
 FT REPEAT 1192 1467
 FT REPEAT 1192 1467
 FT DOMAIN 1 146
 FT TRANSMEM 127 145
 FT DOMAIN 146 163
 FT TRANSMEM 164 183
 FT DOMAIN 184 195
 FT TRANSMEM 196 214
 FT DOMAIN 215 235
 FT TRANSMEM 236 254
 FT DOMAIN 255 273
 FT TRANSMEM 274 293
 FT DOMAIN 294 381
 FT TRANSMEM 382 406
 FT TRANSMEM 382 406

```

FT DOMAIN 407 523 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 524 543 S1 OF REPEAT II (POTENTIAL).
FT DOMAIN 544 558 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 559 577 S2 OF REPEAT II (POTENTIAL).
FT DOMAIN 578 585 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 586 604 S3 OF REPEAT II (POTENTIAL).
FT DOMAIN 605 614 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 615 633 S4 OF REPEAT II (POTENTIAL).
FT DOMAIN 634 652 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 653 673 S5 OF REPEAT II (POTENTIAL).
FT DOMAIN 674 727 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 728 752 S6 OF REPEAT II (POTENTIAL).
FT DOMAIN 753 886 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 887 905 S1 OF REPEAT III (POTENTIAL).
FT DOMAIN 906 921 S2 OF REPEAT III (POTENTIAL).
FT TRANSMEM 922 941 S3 OF REPEAT III (POTENTIAL).
FT DOMAIN 942 953 S4 OF REPEAT III (POTENTIAL).
FT TRANSMEM 954 972 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 973 998 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 999 1017 S5 OF REPEAT III (POTENTIAL).
FT DOMAIN 1018 1037 S6 OF REPEAT III (POTENTIAL).
FT TRANSMEM 1038 1127 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1128 1148 S7 OF REPEAT III (POTENTIAL).
FT TRANSMEM 1149 1205 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1206 1224 S1 OF REPEAT IV (POTENTIAL).
FT TRANSMEM 1225 1239 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1240 1259 S2 OF REPEAT IV (POTENTIAL).
FT TRANSMEM 1260 1266 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1267 1288 S3 OF REPEAT IV (POTENTIAL).
FT TRANSMEM 1289 1313 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1314 1333 S4 OF REPEAT IV (POTENTIAL).

Query Match 79.4%; Score 27; DB 1; Length 2161;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWMWNR 6
Db 3 MWMWNR 8

RESULT 5
CCAD RAT STANDARD; PRT: 2203 RA.
AC P27732; 009022; 009023; 001542; Q62691; Q62815; Q63491;
AC Q63492;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Voltage-dependent L-type calcium channel alpha-1D subunit (Calcium
DE channel, L type, alpha-1 polypeptide, isoform 2) (RAT brain class D)
DE (RBD).
GN CACNA1D OR CACNA1A2 OR CCH1A2 OR CACH3 OR CACNA4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Insulinoma;
RX MEDLINE=95280950; PubMed=7760845;
RX Ihara Y., Yamada Y., Fujii Y., Gono T., Yano H., Yasuda K.,
RX Inagaki N., Seino Y., Seino S.;
RT "Molecular diversity and functional characterization of voltage-
RT dependent calcium channels (CACNA4) expressed in pancreatic beta-
RT cells.";
RL Mol. Endocrinol. 9:121-130(1995).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 3; 6; 7; 8 AND 14).
RC TISSUE=Brain;
RX MEDLINE=91299338; PubMed=1648940;
RX Hui A., Ellinor P.T., Krizanova O., Wang J.-J., Diebold R.J.,

```

```

RA Schwartz A.;
RT "Molecular cloning of multiple subtypes of a novel rat brain isoform
RT of the alpha-1 subunit of the voltage-dependent calcium channel.";
RL Neuron 7:35-44(1991).
RN [3]
RP SEQUENCE OF 1-125 FROM N.A.
RX MEDLINE=96040125; PubMed=7553731;
RA Kamp T.J., Mitas M., Fields K.L., Asch S., Chin H., Marban E.,
RA Nirenberg M.;
RT "Transcriptional regulation of the neuronal L-type calcium channel
RT alpha 1D subunit gene.";
RL Cell. Mol. Neurobiol. 15:307-326(1995).
RN [4]
RP SEQUENCE OF 1100-1410 FROM N.A. (ISOFORMS 11 AND 12).
RX TISSUE=Kidney;
RX MEDLINE=93066265; PubMed=1279681;
RA Yu A.S.L., Hebert S.C., Brenner B.M., Lytton J.;
RT "Molecular characterization and nephron distribution of a family of
RT transcripts encoding the pore-forming subunit of Ca2+ channels in the
RT kidney.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10494-10498(1992).
RN [5]
RP SEQUENCE OF 1218-1498 FROM N.A. (ISOFORM 13).
RX TISSUE=Osteosarcoma;
RX MEDLINE=96074617; PubMed=7479909;
RA Barry E.L.R., Gesek F.A., Froehner S.C., Friedman P.A.;
RT "Multiple calcium channel transcripts in rat osteosarcoma cells:
RT selective activation of alpha 1D isoform by parathyroid hormone.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:10914-10918(1995).
RN [6]
RP SEQUENCE OF 1200-1493 FROM N.A. (ISOFORMS 3; 4; 5; 9 AND 10).
RX TISSUE=Hepatoma;
RX MEDLINE=97376179; PubMed=9232351;
RA Brereton H.M., Harland M.L., Frosio M., Petronijevic T.,
RA Barritt G.J.;
RT "Novel variants of voltage-operated calcium channel alpha-1 subunit
RT transcripts in a rat liver-derived cell line: deletion in the IVS4
RT voltage sensing region.";
RL Cell Calcium 22:39-52(1997).
RN [7]
RP SEQUENCE OF 1307-1479 FROM N.A. (ISOFORM 3).
RX MEDLINE=90239020; PubMed=1692134;
RA Stutch T.P., Leonard J.P., Gilbert M.M., Lester H.A., Davidson N.;
RT "Rat brain expresses a heterogeneous family of calcium channels.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3391-3395(1990).
CC -1- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1D
CC GIVES RISE TO L-TYPE CALCIUM CURRENTS. LONG-LASTING (L-TYPE)
CC CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)
CC GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP),
CC PHENALKYLAMINES, BENZOTHAZEPINES, AND BY OMEGA-AGATOXIN-IIIa
CC (OMEGA-AGA-IIIa). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-
CC GVIA (OMEGA-CXV-GVIA) AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA).
CC -1- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=14;
CC Comment=The region sequenced in isoforms ROB3 and RKC5 is
CC identical to CACNA4;
CC Name=1; Synonyms=CACNA4;
CC IsoId=P27732-1; Sequence=Displayed;
CC Name=2; Synonyms=CACNA4B;
CC IsoId=P27732-2; Sequence=VSP_000923, VSP_000924;
CC Name=3; Synonyms=CACH3A, RB48, RBD-55;

```

CC IsoId=P27732-3; Sequence=VSP_000921;
CC Name=4; Synonyms=Delta-IV-S3;
CC IsoId=P27732-4; Sequence=VSP_000919;
CC Name=5; Synonyms=Delta-IV-S4;
CC IsoId=P27732-5; Sequence=VSP_000922;
CC Name=6; Synonyms=RB9;
CC IsoId=P27732-6; Sequence=VSP_000920, VSP_000921;
CC Name=7; Synonyms=RB11;
CC IsoId=P27732-7; Sequence=VSP_000917;
CC Name=8; Synonyms=RB34;
CC IsoId=P27732-8; Sequence=VSP_000916;
CC Name=9; Synonyms=RH1;
CC IsoId=P27732-9; Sequence=VSP_000918;
CC Name=10; Synonyms=RH2;
CC IsoId=P27732-10; Sequence=VSP_000919, VSP_000922;
CC Name=11; Synonyms=RKC5;
CC IsoId=P27732-13; Sequence=Not described;
CC Name=12; Synonyms=RKC6;
CC IsoId=P27732-11; Sequence=VSP_000919;
CC Name=13; Synonyms=ROB3;
CC IsoId=P27732-14; Sequence=Not described;
CC Name=14; Synonyms=Truncated;
CC IsoId=P27732-12; Sequence=VSP_000925, VSP_000926;
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, PANCREATIC ISLETS AND B-LYMPHOCYTES.
CC -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
CC -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; D38101; BAA07282.1; -
CC EMBL; D38102; BAA07283.1; -
CC EMBL; M57682; AAA42015.1; -
CC EMBL; U14005; AAB60515.1; -
CC EMBL; M99221; AAA40895.1; -
CC EMBL; U31772; AAB89156.1; -
CC EMBL; U49126; AAB61634.1; -
CC EMBL; U49127; AAB61635.1; -
CC EMBL; U49128; AAB61636.1; -
CC InterPro; IPR001682; Ca/Na_pore.
CC InterPro; IPR002077; Ca_channel.
CC InterPro; IPR002111; Cat_channel_TrpL.
CC InterPro; IPR005821; Ion_trans.
CC InterPro; IPR003091; K_channel.
CC InterPro; IPR005446; LVPCALphal.
CC InterPro; IPR005452; LVDCALphal.
CC InterPro; IPR005820; M-channel_nlg.
CC Pfam; PF00520; Ion_trans; 4.
CC PRINTS; PR00167; KCHANNEL.
CC PRINTS; PR00169; KCHANNEL.
CC PRINTS; PR01630; LVDCALPHAL.
CC PRINTS; PR01636; LVDCALPHAL.
CC Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW Calcium channel; Glycoprotein; Repeat; Multigene family;
FT Calcium-binding; Phosphorylation; Alternative splicing.
KW REPEAT 112 408
FT REPEAT 528 774
FT REPEAT 892 1174
FT REPEAT 1211 1486
FT DOMAIN 1 126
FT TRANSMEM 127 145
FT DOMAIN 146 163

FT TRANSMEM 164 183
FT DOMAIN 184 195
FT TRANSMEM 196 214
FT DOMAIN 215 235
FT TRANSMEM 236 254
FT DOMAIN 255 273
FT TRANSMEM 274 293
FT DOMAIN 294 381
FT TRANSMEM 382 406
FT DOMAIN 407 582
FT TRANSMEM 583 602
FT DOMAIN 603 617
FT TRANSMEM 618 636
FT DOMAIN 637 644
FT TRANSMEM 645 663
FT DOMAIN 664 673
FT TRANSMEM 674 692
FT DOMAIN 693 711
FT TRANSMEM 712 732
FT DOMAIN 733 786
FT TRANSMEM 787 811
FT DOMAIN 812 945
FT TRANSMEM 946 964
FT DOMAIN 965 980
FT TRANSMEM 981 1000
Query Match 79.4%; Score 27; DB 1; Length 2203;
Best Local Similarity 83.3%; Pred. No. 83;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWMWNR 6
DB 3 MWMWNR 8
RESULT 6
RIF2 YEAST
ID RIF2 YEAST STANDARD; PRT; 395 AA.
AC Q06208;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE RIF2 protein (RAP1-interacting factor 2).
GN RIF2 OR YLR453C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Bacteria; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3288C;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansgorge W., Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A., Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K., Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P., Louis E.J., Messinguy F., Mewes H.-W., Miosga T., Moestl D., Mueller-Auer S., Neutwich U., Obermaier B., Piravandi E., Pohl T.M., Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M., Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S., Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhaesselt P., Vierendeels F., Voet M., Voickaert G., Voss H., Wambutt R., Wedler E., Wedler H., Zimmermann F.K., Zollner A., Hani J., Hobeisel J.D.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RN Nature 387:87-90(1997).
RL [2]
RN FUNCTION AND SUBUNIT.
RP MEDLINE=97242430; PubMed=9087429;
RX Wotton D., Shore D.;
RA "A novel Rap1-interacting factor, Rif2p, cooperates with Rif1p to regulate telomere length in Saccharomyces cerevisiae.";
RT Genes Dev. 11:748-760(1997).
RL -!- FUNCTION: Involved in transcriptional silencing and telomere

length regulation. Its role in telomere length regulation results from either a block in elongation or promoting degradation of the telomere ends. Loss of Rf1 function results in derepression of an HMR silencer, whose ARS consensus element has been deleted, and in the elongation of telomeres. RAP1 may target the binding of Rf1 to silencers and telomeres.

-/- SUBUNIT: Interacts with Rf1 and RAP1 C-terminus.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U22382; AAC67535.1; -
 PIR; S55975;
 SGD; S0004445; RIF2.
 GO; GO:0005696; C:telomere; IDA.
 GO; GO:0042162; F:telomeric DNA binding activity; IDA.
 GO; GO:0007004; P:telomerase-dependent telomere maintenance; IMP.
 Telomere.
 KW Telomere.
 SQ SEQUENCE 395 AA; 45642 MW; FF6E3A23C8805DE9 CRC64;

Query Match 76.5%; Score 26; DB 1; Length 395;
 Best Local Similarity 83.3%; Pred. No. 30;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M R 6
 |||||
 Db 191 M M M M M R 196

RESULT 7

HRPN_ERWAM STANDARD; PRT; 403 AA.

ID HRPN_ERWAM
 AC Q01099;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Harpin (Harpin-EA).
 GN HRPN.
 OS Erwinia amylovora.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Erwinia.
 OX NCBI_TaxID=552;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
 RC STRAIN=EA321.
 RX MEDLINE=9320201; PubMed=1621099;
 RA Wei Z.-W., Laby R.J., Zumoff C.H., Bauer D.W., He S.Y., Collmer A.,
 RA Beer S.V.;
 RT "Harpin, elicitor of the hypersensitive response produced by the
 RT plant pathogen Erwinia amylovora.";
 RL Science 257:85-88(1992).
 RN [2]
 RP REVISIONS.
 RC STRAIN=EA321;
 RA Laby R.J., Kim J.P., Beer S.V.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -/- FUNCTION: ELICITS THE HYPERSENSITIVE RESPONSE (HR) IN THE PLANT
 CC UPON INFECTION. HARPIN ELICITS HR IN NON-HOSTS AND IS ALSO
 CC REQUIRED FOR PATHOGENICITY IN HOST PLANTS.
 CC -/- SUBCELLULAR LOCATION: SECRETED, VIA THE HRP SECRETION PATHWAY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M92994; AAC31644.2; -
 DR PIR; T08471; T08471.
 DR Hypersensitive response.
 KW Glycerol 1 270
 FT DOMAIN 63
 FT POLY-MET.
 SQ SEQUENCE 403 AA; 39697 MW; 146FA642351D8E87 CRC64;

Query Match 76.5%; Score 26; DB 1; Length 403;
 Best Local Similarity 71.4%; Pred. No. 30;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 M M M M M R 7
 |||||
 Db 63 M M M M M R 69

RESULT 8

POPI_CABEL STANDARD; PRT; 438 AA.

ID POPI_CABEL
 AC Q10666;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Pop-1 protein.
 GN POP-1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2.
 RX MEDLINE=96069861; PubMed=7585963;
 RA Lin R., Thompson S., Priess J.R.;
 RT "pop-1 encodes an HMG box protein required for the specification of a
 RT mesoderm precursor in early C. elegans embryos.";
 RL Cell 83:599-609(1995).
 RN [2]
 RP REVISIONS.
 RA Lin R., Thompson S., Priess J.R.;
 RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP FUNCTION, AND INTERACTION WITH POP-1.
 RX PubMed=11742996;
 RA Calvo D., Victor M., Gay F., Sui G., Luke M.P.-S., Dufourcq P.,
 RA Wen G., Maduro M., Rothman J., Shi Y.;
 RT "A POP-1 repressor complex restricts inappropriate cell type-specific
 RT gene transcription during Caenorhabditis elegans embryogenesis.";
 RL EMBO J. 20:7197-7208(2001).
 CC -/- FUNCTION: Essential for the specification of the mesodermal cell
 CC fate in early embryos. Represses expression of target genes via
 CC its interaction with hda-1 histone deacetylase.
 CC -/- SUBUNIT: Interacts with hda-1.
 CC -/- SUBCELLULAR LOCATION: Nuclear.
 CC -/- DEVELOPMENTAL STAGE: Expressed maternally and zygotically.
 CC -/- SIMILARITY: Contains 1 HMG box domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; U37532; AAC05308.1; -
 DR HSP; P27782; 2LEF.
 DR TRANSFAC; T03244; -
 DR InterPro; IPR000910; HMG_12_box.
 DR Pfam; PF00505; HMG_box; 1.
 DR SMART; SM00398; HMG; 1.
 DR PROSITE; PS00118; HMG_BOX_2; 1.

```

KW DNA-binding; Nuclear protein; Developmental protein.
FT DOMAIN 55 173 PRO-RICH.
FT DOMAIN 132 139 POLY-ALA.
FT DNA_BIND 192 262 RMG_BOX.
FT DOMAIN 340 347 POLY-SER.
FT DOMAIN 359 362 POLY-GLN.
FT DOMAIN 411 416 POLY-GLU.
SQ SEQUENCE 438 AA; 48595 MW; F94073BE40B4095F CRC64;

Query Match 76.5%; Score 26; DB 1; Length 438;
Best Local Similarity 71.4%; Pred. No. 32;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAMMMRL 7
Db 369 MLMQMRL 375

RESULT 9
AMD2_XENLA * STANDARD; PRT; 875 AA.
AC P12890;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Peptidyl-glycine alpha-amidating monooxygenase II precursor
DE (EC 1.14.17.3) (Peptide C-terminal alpha-amidating enzyme II) (AE-II).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=88134244; PubMed=2829895;
RA Ohsuye K., Kitano K., Wada Y., Fuchimura K., Tanaka S., Mizuno K.,
RA Matsuo H.;
RT Cloning of cDNA encoding a new peptide C-terminal alpha-amidating
RT enzyme having a putative membrane-spanning domain from Xenopus laevis
RT skin.
RL Biochem. Biophys. Res. Commun. 150:1275-1281(1988).
CC -1- FUNCTION: C-terminal alpha-amidation of peptides. The reaction
CC produces a peptidyl(2-hydroxyglycine) intermediate is unstable and
CC dismutates to glyoxylate and the corresponding desglycine peptide
CC amide.
CC -1- CATALYTIC ACTIVITY: Peptidylglycine + ascorbate + O(2) =
CC Peptidyl(2-hydroxyglycine) + dehydroascorbate + H(2)O.
CC -1- COFACTOR: COPPER AND ASCORBATE.
CC -1- SIMILARITY: BELONGS TO THE COPPER TYPE II, ASCORBATE-DEPENDENT
CC MONOOXYGENASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; M19032; AAA49667.1;
FIR; A2715; UFXLA2.
HSSP; P14925; 1PHM.
InterPro; IPR000323; Cu2_monooxygenase.
InterPro; IPR001258; NHL.
InterPro; IPR000720; Pamonoxigenase.
Pfam; PF03712; Cu2_monoox_C; 1.
Pfam; PF01082; Cu2_monooxigen; 1.
Pfam; PF01436; NHL; 4.
PRINTS; PR00790; PAMONOXIGENASE.
PROSITE; PS00084; CU2_MONOOXYGENASE_1; 1.
PROSITE; PS00085; CU2_MONOOXYGENASE_2; 1.
Oxidoreductase; Monooxygenase; Copper; Vitamin C; Transmembrane;

KW DNA-binding; Nuclear protein; Signal.
FT SIGNAL 1 39
FT CHAIN 40 875
FT DOMAIN 40 763
FT TRANSXEM 764 787
FT DOMAIN 788 875
FT TRANSXEM 876 890
FT DOMAIN 891 904
FT DOMAIN 905 918
FT DOMAIN 919 932
FT DOMAIN 933 946
FT DOMAIN 947 960
FT DOMAIN 961 974
FT DOMAIN 975 988
FT DOMAIN 989 1002
FT DOMAIN 1003 1016
FT DOMAIN 1017 1030
FT DOMAIN 1031 1044
FT DOMAIN 1045 1058
FT DOMAIN 1059 1072
FT DOMAIN 1073 1086
FT DOMAIN 1087 1100
FT DOMAIN 1101 1114
FT DOMAIN 1115 1128
FT DOMAIN 1129 1142
FT DOMAIN 1143 1156
FT DOMAIN 1157 1170
FT DOMAIN 1171 1184
FT DOMAIN 1185 1198
FT DOMAIN 1199 1212
FT DOMAIN 1213 1226
FT DOMAIN 1227 1240
FT DOMAIN 1241 1254
FT DOMAIN 1255 1268
FT DOMAIN 1269 1282
FT DOMAIN 1283 1296
FT DOMAIN 1297 1310
FT DOMAIN 1311 1324
FT DOMAIN 1325 1338
FT DOMAIN 1339 1352
FT DOMAIN 1353 1366
FT DOMAIN 1367 1380
FT DOMAIN 1381 1394
FT DOMAIN 1395 1408
FT DOMAIN 1409 1422
FT DOMAIN 1423 1436
FT DOMAIN 1437 1450
FT DOMAIN 1451 1464
FT DOMAIN 1465 1478
FT DOMAIN 1479 1492
FT DOMAIN 1493 1506
FT DOMAIN 1507 1520
FT DOMAIN 1521 1534
FT DOMAIN 1535 1548
FT DOMAIN 1549 1562
FT DOMAIN 1563 1576
FT DOMAIN 1577 1590
FT DOMAIN 1591 1604
FT DOMAIN 1605 1618
FT DOMAIN 1619 1632
FT DOMAIN 1633 1646
FT DOMAIN 1647 1660
FT DOMAIN 1661 1674
FT DOMAIN 1675 1688
FT DOMAIN 1689 1702
FT DOMAIN 1703 1716
FT DOMAIN 1717 1730
FT DOMAIN 1731 1744
FT DOMAIN 1745 1758
FT DOMAIN 1759 1772
FT DOMAIN 1773 1786
FT DOMAIN 1787 1800
FT DOMAIN 1801 1814
FT DOMAIN 1815 1828
FT DOMAIN 1829 1842
FT DOMAIN 1843 1856
FT DOMAIN 1857 1870
FT DOMAIN 1871 1884
FT DOMAIN 1885 1898
FT DOMAIN 1899 1912
FT DOMAIN 1913 1926
FT DOMAIN 1927 1940
FT DOMAIN 1941 1954
FT DOMAIN 1955 1968
FT DOMAIN 1969 1982
FT DOMAIN 1983 1996
FT DOMAIN 1997 2010
FT DOMAIN 2011 2024
FT DOMAIN 2025 2038
FT DOMAIN 2039 2052
FT DOMAIN 2053 2066
FT DOMAIN 2067 2080
FT DOMAIN 2081 2094
FT DOMAIN 2095 2108
FT DOMAIN 2109 2122
FT DOMAIN 2123 2136
FT DOMAIN 2137 2150
FT DOMAIN 2151 2164
FT DOMAIN 2165 2178
FT DOMAIN 2179 2192
FT DOMAIN 2193 2206
FT DOMAIN 2207 2220
FT DOMAIN 2221 2234
FT DOMAIN 2235 2248
FT DOMAIN 2249 2262
FT DOMAIN 2263 2276
FT DOMAIN 2277 2290
FT DOMAIN 2291 2304
FT DOMAIN 2305 2318
FT DOMAIN 2319 2332
FT DOMAIN 2333 2346
FT DOMAIN 2347 2360
FT DOMAIN 2361 2374
FT DOMAIN 2375 2388
FT DOMAIN 2389 2402
FT DOMAIN 2403 2416
FT DOMAIN 2417 2430
FT DOMAIN 2431 2444
FT DOMAIN 2445 2458
FT DOMAIN 2459 2472
FT DOMAIN 2473 2486
FT DOMAIN 2487 2500
FT DOMAIN 2501 2514
FT DOMAIN 2515 2528
FT DOMAIN 2529 2542
FT DOMAIN 2543 2556
FT DOMAIN 2557 2570
FT DOMAIN 2571 2584
FT DOMAIN 2585 2598
FT DOMAIN 2599 2612
FT DOMAIN 2613 2626
FT DOMAIN 2627 2640
FT DOMAIN 2641 2654
FT DOMAIN 2655 2668
FT DOMAIN 2669 2682
FT DOMAIN 2683 2696
FT DOMAIN 2697 2710
FT DOMAIN 2711 2724
FT DOMAIN 2725 2738
FT DOMAIN 2739 2752
FT DOMAIN 2753 2766
FT DOMAIN 2767 2780
FT DOMAIN 2781 2794
FT DOMAIN 2795 2808
FT DOMAIN 2809 2822
FT DOMAIN 2823 2836
FT DOMAIN 2837 2850
FT DOMAIN 2851 2864
FT DOMAIN 2865 2878
FT DOMAIN 2879 2892
FT DOMAIN 2893 2906
FT DOMAIN 2907 2920
FT DOMAIN 2921 2934
FT DOMAIN 2935 2948
FT DOMAIN 2949 2962
FT DOMAIN 2963 2976
FT DOMAIN 2977 2990
FT DOMAIN 2991 3004
FT DOMAIN 3005 3018
FT DOMAIN 3019 3032
FT DOMAIN 3033 3046
FT DOMAIN 3047 3060
FT DOMAIN 3061 3074
FT DOMAIN 3075 3088
FT DOMAIN 3089 3102
FT DOMAIN 3103 3116
FT DOMAIN 3117 3130
FT DOMAIN 3131 3144
FT DOMAIN 3145 3158
FT DOMAIN 3159 3172
FT DOMAIN 3173 3186
FT DOMAIN 3187 3200
FT DOMAIN 3201 3214
FT DOMAIN 3215 3228
FT DOMAIN 3229 3242
FT DOMAIN 3243 3256
FT DOMAIN 3257 3270
FT DOMAIN 3271 3284
FT DOMAIN 3285 3298
FT DOMAIN 3299 3312
FT DOMAIN 3313 3326
FT DOMAIN 3327 3340
FT DOMAIN 3341 3354
FT DOMAIN 3355 3368
FT DOMAIN 3369 3382
FT DOMAIN 3383 3396
FT DOMAIN 3397 3410
FT DOMAIN 3411 3424
FT DOMAIN 3425 3438
FT DOMAIN 3439 3452
FT DOMAIN 3453 3466
FT DOMAIN 3467 3480
FT DOMAIN 3481 3494
FT DOMAIN 3495 3508
FT DOMAIN 3509 3522
FT DOMAIN 3523 3536
FT DOMAIN 3537 3550
FT DOMAIN 3551 3564
FT DOMAIN 3565 3578
FT DOMAIN 3579 3592
FT DOMAIN 3593 3606
FT DOMAIN 3607 3620
FT DOMAIN 3621 3634
FT DOMAIN 3635 3648
FT DOMAIN 3649 3662
FT DOMAIN 3663 3676
FT DOMAIN 3677 3690
FT DOMAIN 3691 3704
FT DOMAIN 3705 3718
FT DOMAIN 3719 3732
FT DOMAIN 3733 3746
FT DOMAIN 3747 3760
FT DOMAIN 3761 3774
FT DOMAIN 3775 3788
FT DOMAIN 3789 3802
FT DOMAIN 3803 3816
FT DOMAIN 3817 3830
FT DOMAIN 3831 3844
FT DOMAIN 3845 3858
FT DOMAIN 3859 3872
FT DOMAIN 3873 3886
FT DOMAIN 3887 3900
FT DOMAIN 3901 3914
FT DOMAIN 3915 3928
FT DOMAIN 3929 3942
FT DOMAIN 3943 3956
FT DOMAIN 3957 3970
FT DOMAIN 3971 3984
FT DOMAIN 3985 3998
FT DOMAIN 3999 4012
FT DOMAIN 4013 4026
FT DOMAIN 4027 4040
FT DOMAIN 4041 4054
FT DOMAIN 4055 4068
FT DOMAIN 4069 4082
FT DOMAIN 4083 4096
FT DOMAIN 4097 4110
FT DOMAIN 4111 4124
FT DOMAIN 4125 4138
FT DOMAIN 4139 4152
FT DOMAIN 4153 4166
FT DOMAIN 4167 4180
FT DOMAIN 4181 4194
FT DOMAIN 4195 4208
FT DOMAIN 4209 4222
FT DOMAIN 4223 4236
FT DOMAIN 4237 4250
FT DOMAIN 4251 4264
FT DOMAIN 4265 4278
FT DOMAIN 4279 4292
FT DOMAIN 4293 4306
FT DOMAIN 4307 4320
FT DOMAIN 4321 4334
FT DOMAIN 4335 4348
FT DOMAIN 4349 4362
FT DOMAIN 4363 4376
FT DOMAIN 4377 4390
FT DOMAIN 4391 4404
FT DOMAIN 4405 4418
FT DOMAIN 4419 4432
FT DOMAIN 4433 4446
FT DOMAIN 4447 4460
FT DOMAIN 4461 4474
FT DOMAIN 4475 4488
FT DOMAIN 4489 4502
FT DOMAIN 4503 4516
FT DOMAIN 4517 4530
FT DOMAIN 4531 4544
FT DOMAIN 4545 4558
FT DOMAIN 4559 4572
FT DOMAIN 4573 4586
FT DOMAIN 4587 4600
FT DOMAIN 4601 4614
FT DOMAIN 4615 4628
FT DOMAIN 4629 4642
FT DOMAIN 4643 4656
FT DOMAIN 4657 4670
FT DOMAIN 4671 4684
FT DOMAIN 4685 4698
FT DOMAIN 4699 4712
FT DOMAIN 4713 4726
FT DOMAIN 4727 4740
FT DOMAIN 4741 4754
FT DOMAIN 4755 4768
FT DOMAIN 4769 4782
FT DOMAIN 4783 4796
FT DOMAIN 4797 4810
FT DOMAIN 4811 4824
FT DOMAIN 4825 4838
FT DOMAIN 4839 4852
FT DOMAIN 4853 4866
FT DOMAIN 4867 4880
FT DOMAIN 4881 4894
FT DOMAIN 4895 4908
FT DOMAIN 4909 4922
FT DOMAIN 4923 4936
FT DOMAIN 4937 4950
FT DOMAIN 4951 4964
FT DOMAIN 4965 4978
FT DOMAIN 4979 4992
FT DOMAIN 4993 5006
FT DOMAIN 5007 5020
FT DOMAIN 5021 5034
FT DOMAIN 5035 5048
FT DOMAIN 5049 5062
FT DOMAIN 5063 5076
FT DOMAIN 5077 5090
FT DOMAIN 5091 5104
FT DOMAIN 5105 5118
FT DOMAIN 5119 5132
FT DOMAIN 5133 5146
FT DOMAIN 5147 5160
FT DOMAIN 5161 5174
FT DOMAIN 5175 5188
FT DOMAIN 5189 5202
FT DOMAIN 5203 5216
FT DOMAIN 5217 5230
FT DOMAIN 5231 5244
FT DOMAIN 5245 5258
FT DOMAIN 5259 5272
FT DOMAIN 5273 5286
FT DOMAIN 5287 5300
FT DOMAIN 5301 5314
FT DOMAIN 5315 5328
FT DOMAIN 5329 5342
FT DOMAIN 5343 5356
FT DOMAIN 5357 5370
FT DOMAIN 5371 5384
FT DOMAIN 5385 5398
FT DOMAIN 5399 5412
FT DOMAIN 5413 5426
FT DOMAIN 5427 5440
FT DOMAIN 5441 5454
FT DOMAIN 5455 5468
FT DOMAIN 5469 5482
FT DOMAIN 5483 5496
FT DOMAIN 5497 5510
FT DOMAIN 5511 5524
FT DOMAIN 5525 5538
FT DOMAIN 5539 5552
FT DOMAIN 5553 5566
FT DOMAIN 5567 5580
FT DOMAIN 5581 5594
FT DOMAIN 5595 5608
FT DOMAIN 5609 5622
FT DOMAIN 5623 5636
FT DOMAIN 5637 5650
FT DOMAIN 5651 5664
FT DOMAIN 5665 5678
FT DOMAIN 5679 5692
FT DOMAIN 5693 5706
FT DOMAIN 5707 5720
FT DOMAIN 5721 5734
FT DOMAIN 5735 5748
FT DOMAIN 5749 5762
FT DOMAIN 5763 5776
FT DOMAIN 5777 5790
FT DOMAIN 5791 5804
FT DOMAIN 5805 5818
FT DOMAIN 5819 5832
FT DOMAIN 5833 5846
FT DOMAIN 5847 5860
FT DOMAIN 5861 5874
FT DOMAIN 5875 5888
FT DOMAIN 5889 5902
FT DOMAIN 5903 5916
FT DOMAIN 5917 5930
FT DOMAIN 5931 5944
FT DOMAIN 5945 5958
FT DOMAIN 5959 5972
FT DOMAIN 5973 5986
FT DOMAIN 5987 6000
FT DOMAIN 6001 6014
FT DOMAIN 6015 6028
FT DOMAIN 6029 6042
FT DOMAIN 6043 6056
FT DOMAIN 6057 6070
FT DOMAIN 6071 6084
FT DOMAIN 6085 6098
FT DOMAIN 6099 6112
FT DOMAIN 6113 6126
FT DOMAIN 6127 6140
FT DOMAIN 6141 6154
FT DOMAIN 6155 6168
FT DOMAIN 6169 6182
FT DOMAIN 6183 6196
FT DOMAIN 6197 6210
FT DOMAIN 6211 6224
FT DOMAIN 6225 6238
FT DOMAIN 6239 6252
FT DOMAIN 6253 6266
FT DOMAIN 6267 6280
FT DOMAIN 6281 6294
FT DOMAIN 6295 6308
FT DOMAIN 6309 6322
FT DOMAIN 6323 6336
FT DOMAIN 6337 6350
FT DOMAIN 6351 6364
FT DOMAIN 6365 6378
FT DOMAIN 6379 6392
FT DOMAIN 6393 6406
FT DOMAIN 6407 6420
FT DOMAIN 6421 6434
FT DOMAIN 6435 6448
FT DOMAIN 6449 6462
FT DOMAIN 6463 6476
FT DOMAIN 6477 6490
FT DOMAIN 6491 6504
FT DOMAIN 6505 6518
FT DOMAIN 6519 6532
FT DOMAIN 6533 6546
FT DOMAIN 6547 6560
FT DOMAIN 6561 6574
FT DOMAIN 6575 6588
FT DOMAIN 6589 6602
FT DOMAIN 6603 6616
FT DOMAIN 6617 6630
FT DOMAIN 6631 6644
FT DOMAIN 6645 6658
FT DOMAIN 6659 6672
FT DOMAIN 6673 6686
FT DOMAIN 6687 6700
FT DOMAIN 6701 6714
FT DOMAIN 6715 6728
FT DOMAIN 6729 6742
FT DOMAIN 6743 6756
FT DOMAIN 6757 6770
FT DOMAIN 6771 6784
FT DOMAIN 6785 6798
FT DOMAIN 6799 6812
FT DOMAIN 6813 6826
FT DOMAIN 6827 6840
FT DOMAIN 6841 6854
FT DOMAIN 6855 6868
FT DOMAIN 6869 6882
FT DOMAIN 6883 6896
FT DOMAIN 6897 6910
FT DOMAIN 6911 6924
FT DOMAIN 6925 6938
FT DOMAIN 6939 6952
FT DOMAIN 6953 6966
FT DOMAIN 6967 6980
FT DOMAIN 6981 6994
FT DOMAIN 6995 7008
FT DOMAIN 7009 7022
FT DOMAIN 7023 7036
FT DOMAIN 7037 7050
FT DOMAIN 7051 7064
FT DOMAIN 7065 7078
FT DOMAIN 7079 7092
FT DOMAIN 7093 7106
FT DOMAIN 7107 7120
FT DOMAIN 7121 7134
FT DOMAIN 7135 7148
FT DOMAIN 7149 7162
FT DOMAIN 7163 7176
FT DOMAIN 7177 7190
FT DOMAIN 7191 7204
FT DOMAIN 7205 7218
FT DOMAIN 7219 7232
FT DOMAIN 7233 7246
FT DOMAIN 7247 7260
FT DOMAIN 7261 7274
FT DOMAIN 7275 7288
FT DOMAIN 7289 7302
FT DOMAIN 7303 7316
FT DOMAIN 7317 7330
FT DOMAIN 7331 7344
FT DOMAIN 7345 7358
FT DOMAIN 7359 7372
FT DOMAIN 7373 7386
FT DOMAIN 7387 7400
FT DOMAIN 7401 7414
FT DOMAIN 7415 7428
FT DOMAIN 7429 7442
FT DOMAIN 7443 7456
FT DOMAIN 7457 7470
FT DOMAIN 7471 7484
FT DOMAIN 7485 7498
FT DOMAIN 7499 7512
FT DOMAIN 7513 7526
FT DOMAIN 7527 7540
FT DOMAIN 7541 7554
FT DOMAIN 7555 7568
FT DOMAIN 7569 7582
FT DOMAIN 7583 7596
FT DOMAIN 7597 7610
FT DOMAIN 7611 7624
FT DOMAIN 7625 7638
FT DOMAIN 7639 7652
FT DOMAIN 7653 7666
FT DOMAIN 7667 7680
FT DOMAIN 7681 7694
FT DOMAIN 7695 7708
FT DOMAIN 7709 7722
FT DOMAIN 7723 7736
FT DOMAIN 7737 7750
FT DOMAIN 7751 7764
FT DOMAIN 7765 7778
FT DOMAIN 7779 7792
FT DOMAIN 7793 7806
FT DOMAIN 7807 7820
FT DOMAIN 7821 7834
FT DOMAIN 7835 7848
FT DOMAIN 7849 7862
FT DOMAIN 7863 7876
FT DOMAIN 7877 7890
FT DOMAIN 7891 7904
FT DOMAIN 7905 7918
FT DOMAIN 7919 7932
FT DOMAIN 7933 7946
FT DOMAIN 7947 7960
FT DOMAIN 7961 7974
FT DOMAIN 7975 7988
FT DOMAIN 7989 8002
FT DOMAIN 8003 8016
FT DOMAIN 8017 8030
FT DOMAIN 8031 8044
FT DOMAIN 8045 8058
FT DOMAIN 8059 8072
FT DOMAIN 8073 8086
FT DOMAIN 8087 8100
FT DOMAIN 8101 8114
FT DOMAIN 8115 8128
FT DOMAIN 8129 8142
FT DOMAIN 8143 8156
FT DOMAIN 8157 8170
FT DOMAIN 8171 8184
FT DOMAIN 8185 8198
FT DOMAIN 8199 8212
FT DOMAIN 8213 8226
FT DOMAIN 8227 8240
FT DOMAIN 8241 8254
FT DOMAIN 8255 8268
FT DOMAIN 8269 8282
FT DOMAIN 8283 8296
FT DOMAIN 8297 8310
FT DOMAIN 8311 8324
FT DOMAIN 8325 8338
FT DOMAIN 8339 8352
FT DOMAIN 8353 8366
FT DOMAIN 8367 8380
FT DOMAIN 8381 8394
FT DOMAIN 8395 8408
FT DOMAIN 8409 8422
FT DOMAIN 8423 8436
FT DOMAIN 8437 8450
FT DOMAIN 8451 8464
FT DOMAIN 8465 8478
FT DOMAIN 8479 8492
FT DOMAIN 8493 8506
FT DOMAIN 8507 8520
FT DOMAIN 8521 8534
FT DOMAIN 8535 8548
FT DOMAIN 8549 8562
FT DOMAIN 8563 8576
FT DOMAIN 8577 8590
FT DOMAIN 8591 8604
FT DOMAIN 8605 8618
FT DOMAIN 8619 8632
FT DOMAIN 8633 8646
FT DOMAIN 8647 8660
FT DOMAIN 8661 8674
FT DOMAIN 8675 8688
FT DOMAIN 8689 8702
FT DOMAIN 8703 8716
FT DOMAIN 8717 8730
FT DOMAIN 8731 8744
FT DOMAIN 8745 8758
FT DOMAIN 8759 8772
FT DOMAIN 8773 8786
FT DOMAIN 8787 8800
FT DOMAIN 8801 8814
FT DOMAIN 8815 8828
FT DOMAIN 8829 8842
FT DOMAIN 8843 8856
FT DOMAIN 8857 8870
FT DOMAIN 8871 8884
FT DOMAIN 8885 8898
FT DOMAIN 8899 8912
FT DOMAIN 8913 8926
FT DOMAIN 8927 8940
FT DOMAIN 8941 8954
FT DOMAIN 8955 8968
FT DOMAIN 8969 8982
FT DOMAIN 8983 8996
FT DOMAIN 8997 9010
FT DOMAIN 9011 9024
FT DOMAIN 9025 9038
FT DOMAIN 9039 9052
FT DOMAIN 9053 9066
FT DOMAIN 9067 9080
FT DOMAIN 9081 9094
FT DOMAIN 9095 9108
FT DOMAIN 9109 9122
FT DOMAIN 9123 9136
FT DOMAIN 9137 9150
FT DOMAIN 9151 9164
FT DOMAIN 9165 9178
FT DOMAIN 9179 9192
FT DOMAIN 9193 9206
FT DOMAIN 9207 9220
FT DOMAIN 9221 9234
FT DOMAIN 9235 9248
FT DOMAIN 9249 9262
FT DOMAIN 9263 9276
FT DOMAIN 9277 9290
FT DOMAIN 9291 9304
FT DOMAIN 9305 9318
FT DOMAIN 9319 9332
FT DOMAIN 9333 9346
FT DOMAIN 9347 9360
FT DOMAIN 9361 9374
FT DOMAIN 9375 9388
FT DOMAIN 9389 9402
FT DOMAIN 9403 9416
FT DOMAIN 9417 9430
FT DOMAIN 9431 9444
FT DOMAIN 9445 9458
FT DOMAIN 9459 9472
FT DOMAIN 9473 9486
FT DOMAIN 9487 9500
FT DOMAIN 9501 9514
FT DOMAIN 9515 9528
FT DOMAIN 9529 9542
FT DOMAIN 9543 9556
FT DOMAIN 9557 9570
FT DOMAIN 9571 9584
FT DOMAIN 9585 9598
FT DOMAIN 9599 9612
FT DOMAIN 9613 9626
FT DOMAIN 9627 9640
FT DOMAIN 9641 9654
FT DOMAIN 9655 9668
FT DOMAIN 9669 9682
FT DOMAIN 9683 9696
FT DOMAIN 9697 9710
FT DOMAIN 9711 9724
FT DOMAIN 9725 9738
FT DOMAIN 9739 9752
FT DOMAIN 9753 9766
FT DOMAIN 9767 9780
FT DOMAIN 9781 9794
FT DOMAIN 9795 9808
FT DOMAIN 9809 9822
FT DOMAIN 9823 9836
FT DOMAIN 9837 9850
FT DOMAIN 9851 9864
FT DOMAIN 9865 9878
FT DOMAIN 9879 9892
FT DOMAIN 9893 9906
FT DOMAIN 9907 9920
FT DOMAIN 9921 9934
FT DOMAIN 9935 9948
FT DOMAIN 9949 9962
FT DOMAIN 9963 9976
FT DOMAIN 9977 9990
FT DOMAIN 9991 10000

Query Match 73.5%; Score 25; DB 1; Length 313;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMMMRL 5
Db 83 MAMMMRL 87

```

RESULT 11
ID AOX1 SOYBN STANDARD; PRT; 321 AA.
AC Q07185; Q41265;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alternative oxidase 1, mitochondrial precursor (EC 1.-.-.-).
GN AOX1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Shoot;
RX MEDLINE=94120028; PubMed=8290651;
RA Whelan J.M., McIntosh L., Day D.A.;
RT "Sequencing of a soybean alternative oxidase cDNA clone.";
RL Plant Physiol. 103:1481-1481(1993).
RN [2]
RN SEQUENCE OF 194-233 FROM N.A.
RX MEDLINE=96165778; PubMed=8580775;
RA Whelan J., Millar A.H., Day D.A.;
RT "The alternative oxidase is encoded in a multigene family in soybean.";
RL Planta 199:197-201(1996).
CC -!- FUNCTION: CATALYZES CYANIDE-RESISTANT OXYGEN CONSUMPTION. MAY INCREASE RESPIRATION WHEN THE CYTOCHROME RESPIRATORY PATHWAY IS RESTRICTED, OR IN RESPONSE TO LOW TEMPERATURES (BY SIMILARITY).
CC -!- PATHWAY: Alternative respiratory pathway.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (PROBABLE).
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL, POSSIBLY IN THE INNER SURFACE OF THE INNER MITOCHONDRIAL MEMBRANE.
CC -!- INDUCTION: By salicylic acid.
CC -!- SIMILARITY: BELONGS TO THE ALTERNATIVE OXIDASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
DR EMBL; X68702; CAA48653.1; -;
DR EMBL; S81466; AAC34192.1; -;
DR InterPro; IPR002880; AOX.
KW Oxidoreductase; Transmembrane; Mitochondrion; Respiratory chain; Inner membrane; Transmembrane; Multigene family.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? 321 ALTERNATIVE OXIDASE 1.
FT TRANSMEM 146 166 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
FT DISULFID 94 94 INTERCHAIN (POTENTIAL).
FT METAL 112 112 POTENTIAL.
FT METAL 165 165 POTENTIAL.
FT METAL 192 192 POTENTIAL.
FT METAL 233 233 POTENTIAL.
FT METAL 294 294 POTENTIAL.
FT METAL 299 299 POTENTIAL.
FT DOMAIN 1 5 POLY-MET.
SQ SEQUENCE 321 AA; 36437 MW; E0FF981F0C2D0E7 CRC64;
Query Match 73.5%; Score 25; DB 1; Length 321;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 M00000 5
|||||

DB 1 M00000 5
RESULT 12
ID CY1 HUMAN STANDARD; PRT; 325 AA.
AC P08574;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cytochrome c1, heme protein, mitochondrial precursor (Cytochrome c-1).
GN CYC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=89109139; PubMed=2536365;
RA Suzuki H., Hosokawa Y., Nishikimi M., Ozawa T.;
RT "Structural organization of the human mitochondrial cytochrome c1 gene.";
RL J. Biol. Chem. 264:1368-1374(1989).
RN [2]
RN SEQUENCE FROM N.A., AND VARIANT VAL-89.
RX MEDLINE=88233946; PubMed=2836796;
RA Nishikimi M., Ohta S., Suzuki H., Tanaka T., Kikkawa F., Tanaka M., Kagawa Y., Ozawa T.;
RT "Nucleotide sequence of a cDNA encoding the precursor to human cytochrome c1.";
RL Nucleic Acids Res. 16:3577-3577(1988).
RN [3]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain, Lung, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Scheraga H., Shenmen C.M., Schuler G.D., Klausner R.D., Collins F.S., Wagner K.H., Scheaffer C.F., Bhat N.K., Altschul S.F., Zeeberg B., Buetow K.H., Wang J., Hsieh F., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L., Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stappleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RN SEQUENCE OF 99-325 FROM N.A.
RX MEDLINE=87241521; PubMed=3036122;
RA Nishikimi M., Suzuki H., Ohta S., Sakurai T., Shimomura Y., Tanaka M., Kagawa Y., Ozawa T.;
RT "Isolation of a cDNA clone for human cytochrome c1 from a lambda gt10 expression library.";
RL Biochem. Biophys. Res. Commun. 145:34-39(1987).
RN [5]
RN VARIANT VAL-89.
RX MEDLINE=99381536; PubMed=10453733;
RA Valnot I., Kassis J., Chretien D., de Lonlay P., Parfait B., Munich A., Kachaner J., Rustin P., Roetig A.;
RT "A mitochondrial cytochrome b mutation but no mutations of nuclearly encoded subunits in ubiquinol cytochrome c reductase (complex III) deficiency.";
RL Hum. Genet. 104:460-466(1999).
CC -!- FUNCTION: THIS IS THE HEME-CONTAINING COMPONENT OF THE CYTOCHROME

B-C1 COMPLEX, WHICH ACCEPTS ELECTRONS FROM RIBSKE PROTEIN AND TRANSFERS ELECTRONS TO CYTOCHROME C IN THE MITOCHONDRIAL RESPIRATORY CHAIN.

-1- SUBUNIT: BCL COMPLEX CONTAINS 11 SUBUNITS; 3 RESPIRATORY SUBUNITS (CYTOCHROME B, CYTOCHROME C1, RIBSKE PROTEIN), 2 CORE PROTEINS AND 6 LOW-MOLECULAR WEIGHT PROTEINS.

-1- SUBCELLULAR LOCATION: Mitochondrial intermembrane space.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; J04444; AAA52135.1; -
 EMBL; M16597; AAA35730.1; -
 EMBL; BC001006; AAH01006.1; -
 EMBL; BC015616; AAH15616.1; -
 EMBL; BC020566; AAH20566.1; -
 EMBL; X06994; CAA30052.1; -
 PIR; A31481; S00680.
 Genew; HGNC:2579; CYC1.
 MIM; 123980; -
 GO; GO:0005739; C:mitochondrion; TAS.
 GO; GO:0009461; F:cytochrome c; TAS.
 InterPro; IPR002326; Cyt C1.
 InterPro; IPR000345; CytC_heme_bind.
 Pfam; PF02167; Cytochrome_C1; 1.
 PRINTS; PS06003; CYTOCHROME_C1.
 PROSITE; PS00190; CYTOCHROME C; 1.
 Electron transport; Respiratory chain; Heme; Mitochondrion;
 Transmembrane; Transit peptide; Polymorphism.
 TRANSIT 85 325 MITOCHONDRION.
 FT CHAIN 1 84 CYTOCHROME C1, HEME PROTEIN.
 BINDING 121 121 HEME (COVALENT).
 BINDING 124 124 HEME (COVALENT).
 METAL 125 125 IRON (HEME AXIAL LIGAND).
 METAL 244 244 IRON (HEME AXIAL LIGAND). (BY SIMILARITY).
 TRANSMEM 292 306 ANCHORS TO THE MEMBRANE (POTENTIAL).
 VARIANT 89 89 L -> V.
 FT SEQUENCE 325 AA; 35390 MW; CC98152608992BDC CRC64;
 Query Match 73.5%; Score 25; DB 1; Length 325;
 Best Local Similarity 71.4%; Pred. NO. 43;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAMMMML 7
 Db 292 MAMMMML 298

RESULT 13
 HKL6_LYCES STANDARD; PRT; 355 AA.
 AC O22299;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein knotted-1 like LFN6.
 GN LFN6.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Lamiales; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. VENT Cherry;
 RX MEDLINE=98145476; Pubmed=949482;
 RA Janssen B.J., Williams A., Chen J.J., Mathern J., Hake S., Sinha N.;

Isolation and characterization of two knotted-like homeobox genes from tomato".
 Plant Mol. Biol. 36:417-425(1998).
 -1- FUNCTION: MAY HAVE A ROLE TO PLAY IN FORMATIVE EVENTS IN OVULE AND EMBRIO MORPHOGENESIS. PROBABLY BINDS TO THE DNA SEQUENCE 5'-TGAC-3'.

-1- SUBCELLULAR LOCATION: Nuclear (Probable).
 -1- TISSUE SPECIFICITY: EXPRESSED IN DEVELOPING LATERAL ORGANS AND DEVELOPING OVARIES IN FLOWERS.
 -1- SIMILARITY: BELONGS TO THE TALE/KNOX HOMEBOX FAMILY.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF000141; AAC49917.1; -
 PIR; T04317; T04317.
 KSSP; P40424; 1B72.
 TRANSFAC; T04057; -
 InterPro; IPR005539; ELK.
 InterPro; IPR001356; Homeobox.
 InterPro; IPR005540; KNOX1.
 InterPro; IPR005541; KNOX2.
 Pfam; PF03789; ELK; 1.
 Pfam; PF00046; homeobox; 1.
 Pfam; PF03790; KNOX1; 1.
 Pfam; PF03791; KNOX2; 1.
 ProDom; PD000010; Homeobox; 1.
 SMART; SM00389; HOX; 1.
 PROSITE; PS00027; HOMEBOX 1; 1.
 PROSITE; PS00071; HOMEBOX 2; 1.
 KW Homeobox; DNA-binding; Nuclear protein.
 FT DOMAIN 24 33 POLY-ASN.
 FT DOMAIN 47 51 POLY-MET.
 FT DOMAIN 52 56 POLY-PRO.
 FT DOMAIN 60 63 POLY-ASN.
 FT DOMAIN 79 84 POLY-ASN.
 FT DOMAIN 91 95 POLY-SER.
 FT DOMAIN 142 146 POLY-SER.
 FT DOMAIN 235 258 ELK DOMAIN.
 FT DNA_BIND 259 321 HOMEBOX (TALE-TYPE).
 SQ SEQUENCE 355 AA; 39796 MW; 9E9A1AFD75808C49 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 355;
 Best Local Similarity 100.0%; Pred. NO. 47;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAMMMML 5
 Db 47 MAMMMML 51

RESULT 14
 Y4WD_RHISN STANDARD; PRT; 377 AA.
 ID Y4WD_RHISN
 AC P55682;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical transport protein Y4WD.
 GN Y4WD.
 OS Rhizobium sp. (strain NGR234).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
 OX NCBI_TaxID=394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97305956; Pubmed=9163424;

RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RA "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -!- FUNCTION: COULD BE INVOLVED IN A TRANSPORT SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: TO R.MELILOTI MOSC.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000103; AAB91911.1; -
DR InterPro: IPR007114; MIPS.
KW Hypothetical protein; Transmembrane; Transport; Plasmid.
FT TRANSMEM 4 24 POTENTIAL.
FT TRANSMEM 41 61 POTENTIAL.
FT TRANSMEM 85 105 POTENTIAL.
FT TRANSMEM 134 154 POTENTIAL.
FT TRANSMEM 159 179 POTENTIAL.
FT TRANSMEM 192 212 POTENTIAL.
FT TRANSMEM 278 298 POTENTIAL.
FT TRANSMEM 301 321 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT TRANSMEM 356 376 POTENTIAL.
SQ SEQUENCE 377 AA; 39051 MW; 49CF6E44AA0D74BD CRC64;
Query Match 73.5%; Score 25; DB 1; Length 377;
Best Local Similarity 83.3%; Pred. No. 50;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 MWMRL 7
Db 250 MWMRL 255
RESULT 15
STM_ARATH STANDARD; PRT; 382 AA.
AC Q38874; Q8RXJ1; Q9MAV3;
DT 15-JUL-1998 (Rel. 36, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Homeobox protein SHOOT MERISTEMLESS.
GN STM OR AT1G62360 OR F2401.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]_TaxID=3702;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=96135134; PubMed=8538741;
RA Long J.A., Moan E.I., Medford J.I., Barton M.K.;
RT "A member of the KNOTTED class of homeodomain proteins encoded by the
RT STM gene of Arabidopsis.";
RL Nature 379:66-69(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egtu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,

RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
RA Maltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D., Walker M.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RN Nature 408:816-820(2000).
RP [3]
RP SEQUENCE OF 57-382 FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGSC).";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RN FUNCTION.
RX MEDLINE=21932426; PubMed=11934861;
RA Byrne M.E., Simorowski J., Martienssen R.A.;
RT "ASYMMETRIC LEAVES1 reveals knox gene redundancy in Arabidopsis.";
RL Development 129:1957-1965(2002).
CC -!- FUNCTION: Required for shoot apical meristem formation during
CC embryogenesis. Negatively regulates ASYMMETRIC LEAVES1 (AS1) and
CC ASYMMETRIC LEAVES2 (AS2 or LBD6). Probably binds to the DNA
CC sequence 5'-TGAC-3'.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: Expressed in all four types of shoot apical
CC meristems (SAM) i.e., in vegetative, axillary, inflorescence and
CC floral.
CC -!- DEVELOPMENTAL STAGE: First expressed in early to mid-globular-
CC stage embryos. In late globular stage, detected as a stripe
CC running medially across the top of the embryo. In heart stage
CC embryo, expression is restricted to a notch between the
CC cotyledons. In seedlings and adult plants found in all shoot
CC apical meristems. In the inflorescence meristem, expression
CC disappears as floral buds are initiated and reappears in the later
CC floral meristem where it is found in the central portion of the
CC developing gynoecium. Also detected in the L1 layer of embryo.
CC -!- SIMILARITY: BELONGS TO THE TALE/KNOX HOMEBOX FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32344; AAC49148.1; -
DR EMBL; AC003113; AAF70849.1; -
DR EMBL; AY080857; AAL87330.1; -
DR PIR; S68456; S68456.
DR PIR; T01446; T01446.
DR HSSP; P41778; IDU6.
DR InterPro: IPR005539; ELK.
DR InterPro: IPR001356; Homeobox.
DR InterPro: IPR005540; KNOX1.
DR InterPro: IPR005541; KNOX2.
DR Pfam; PF03789; ELK; 1.
DR Pfam; PF00046; homeobox; 1.
DR Pfam; PF03790; KNOX1; 1.
DR Pfam; PF03791; KNOX2; 1.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
KW DNA-binding; Homeobox; Nuclear protein; Developmental protein.
FT DOMAIN 260 283 ELK DOMAIN.
FT DNA_BIND 284 346 HOMEBOX (TALE-TYPE).

FT DOMAIN	26	30	POLY-MET.
FT DOMAIN	37	60	HIS-RICH.
FT DOMAIN	83	89	POLY-SER.
FT DOMAIN	113	120	POLY-SER.
FT DOMAIN	161	166	POLY-ALA.
FT CONFLICT	115	115	S -> F (IN REF. 1).
FT CONFLICT	229	233	MISSING (IN REF. 2).
FT CONFLICT	389	389	G -> D (IN REF. 1).
SQ SEQUENCE	382 AA;	42753 MW;	6227D3DE4093E732 CRC64;

Query Match 73.5%; Score 25; DB 1; Length 382;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MMMMM	5
Db	26	MMMMMM	30

Search completed: February 11, 2004, 17:04:17
Job time : 6.16667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-901-187C-12

Perfect score: 34
Sequence: 1 MWMWRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR 76:**

1: pirl:**

2: pirl:**

3: pirl:**

4: pirl:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	94.1	70	2 T48381	hypothetical prote
2	30	88.2	192	2 H96828	hypothetical prote
3	30	88.2	249	2 T08410	hypothetical prote
4	30	88.2	353	2 T07725	hypothetical prote
5	27	79.4	39	2 PC4294	high mobility grou
6	27	79.4	146	2 S14946	2S seed storage pr
7	27	79.4	147	2 JQ1320	high mobility grou
8	27	79.4	427	2 T04869	transforming prote
9	27	79.4	1610	2 A46327	voltage-dependent
10	27	79.4	1646	2 JH0422	voltage-dependent
11	27	79.4	2161	2 JH0564	calcium channel al
12	27	79.4	2181	2 A38198	calcium channel al
13	27	79.4	2203	2 T42742	voltage-dependent
14	26	76.5	216	2 H82695	DNA-3-methyladenin
15	26	76.5	302	2 H96811	protein F3F9.20 [i
16	26	76.5	309	2 E83999	mutants block spor
17	26	76.5	340	2 C87732	protein W03D8.4 [i
18	26	76.5	340	2 T32931	hypothetical prote
19	26	76.5	375	2 T06096	hypothetical prote
20	26	76.5	395	2 S5975	hypothetical prote
21	26	76.5	403	2 T08471	harpin - Erwinia a
22	26	76.5	436	2 E96635	hypothetical prote
23	26	76.5	438	1 A57667	pop-1 protein - Ca
24	26	76.5	484	2 A57667	probable TrkH-like
25	26	76.5	875	1 URX1A2	peptidylglycine mo
26	25	73.5	68	2 T03645	hypothetical prote
27	25	73.5	98	2 A98222	hypothetical prote
28	25	73.5	120	2 A12181	hypothetical prote
29	25	73.5	124	2 F81942	probable regulator

30 25 73.5 124 2 A81163 conserved hypothet
31 25 73.5 135 2 F86239 protein F20B24.4 [i
32 25 73.5 176 2 A86159 hypothetical prote
33 25 73.5 197 2 T26156 hypothetical prote
34 25 73.5 217 2 H96657 conserved hypothet
35 25 73.5 220 2 AD0225 probable amino-aci
36 25 73.5 242 2 B2815 conserved hypothet
37 25 73.5 268 2 T04787 hypothetical prote
38 25 73.5 275 2 T51651 myb-related trans
39 25 73.5 289 2 T29216 hypothetical prote
40 25 73.5 292 2 T09030 hypothetical prote
41 25 73.5 307 2 AD0883 conserved hypothet
42 25 73.5 313 2 T11160 NADH2 dehydrogenas
43 25 73.5 321 2 S31711 alternative respir
44 25 73.5 325 1 S00680 ubiquinol-cytochro
45 25 73.5 327 2 T09687 chitinase (EC 3.2.

ALIGNMENTS

RESULT 1

T48381
hypothetical protein F12B4.330 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T48381
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24492
A:Accession: T48381
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-70 <BEV>
A:Cross-references: EMBL:AL162751
A:Experimental source: cultivar Columbia; BAC clone F12E4
C:Genetics:
A:Map position: 5
A:Note: F12B4.330

Query Match 94.1%; Score 32; DB 2; Length 70;
Best Local Similarity 85.7%; Pred. No. 0.57;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWMWRL 7
Db 3 MWMWRL 9

RESULT 2

H96828
hypothetical protein F19K16.24 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 21-Oct-2002
C:Accession: H96828
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H96828
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <STO>
A:Cross-references: GB:AB005173; NID:96453865; PIDN:AAF09049.1; GSPDB:GN00141
C:Genetics:
A:Gene: F19K16.24

A;Map position: 1
C;Superfamily: stellacyanin

Query Match 88.2%; Score 30; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 4.3; 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWMWMMR 6
DB 2 MWMWMMR 7

RESULT 3

T08410
hypothetical protein F18B3.170 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999

C;Accession: T08410

R;Querier, F.; Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Salanoub
submitted to the Protein Sequence Database, May 1999

A;Reference number: 216409

A;Accession: T08410

A;Molecule type: DNA

A;Residues: 1-249 <QUE>

A;Cross-references: EMBL:AL049862; GSPDB:GNC00061; ATSP:F18B3.170

A;Experimental source: cultivar Columbia; BAC clone F18B3

C;Genetics:

A;Gene: ATSP:F18B3.170

A;Map position: 3

Query Match 88.2%; Score 30; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWMWMMR 6
DB 174 MWMWMMR 179

RESULT 4

T07725
hypothetical protein T23J7.190 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 22-Oct-1999

C;Accession: T07725

R;Querier, F.; Choise, N.; Robert, C.; Brottier, P.; Cattolico, L.; Artig
submitted to the Protein Sequence Database, April 1999

A;Reference number: 215793

A;Accession: T07725

A;Molecule type: DNA

A;Residues: 1-353 <QUE>

A;Cross-references: EMBL:AL049746; GSPDB:GNC00061; ATSP:T23J7.190

A;Experimental source: cultivar Columbia; BAC clone T23J7

C;Genetics:

A;Gene: ATSP:T23J7.190

A;Map position: 3

A;Introns: 28/2; 83/3; 165/3; 262/3

Query Match 88.2%; Score 30; DB 2; Length 353;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWMWMMR 6
DB 122 MWMWMMR 127

RESULT 5

PC4294
high mobility group protein Pf16 - malaria parasite (Plasmodium falciparum) (fragments)

C;Species: Plasmodium falciparum

C;Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jun-2000

C;Accession: PC4294

R;Nambiar, A.; Kandala, J.C.; Dolan, S.A.; Jensen, J.B.; Guntaka, R.V.
Biochem. Biophys. Res. Commun. 234, 101-106, 1997
A;Title: Molecular cloning and characterization of a cDNA for the highly conserved HMG-
A;Reference number: JG5403; MUID:97312528; PMID:9168969

A;Accession: PC4294

A;Molecule type: mRNA

A;Residues: 1-39 <NAM>

C;Comment: This protein is involved in DNA replication, chromatin assembly and transcri
C;Genetics:

A;Gene: Pf16

Query Match 79.4%; Score 27; DB 2; Length 39;
Best Local Similarity 83.3%; Pred. No. 5.2;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWMWMMR 6
DB 21 MWMWMMR 26

RESULT 6

S14946

2S seed storage protein large chain - Brazil nut

N;Alternate names: albumin 2S precursor

C;Species: Bertholletia excelsa (Brazil nut)

C;Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jul-2000

C;Accession: S14946; S14479; S06252; S21640; B25802

R;Gander, E.S.; Holmstrom, K.O.; de Paiva, G.R.; de Castro, L.A.B.; Carneiro, M.; Gros
Plant Mol. Biol. 16, 437-448, 1991

A;Title: Isolation, characterization and expression of a gene coding for a 2S albumin f

A;Reference number: S14946; MUID:91370890; PMID:1840693

A;Accession: S14946

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-146 <GAN>

A;Cross-references: EMBL:X54490; NID:gl7710; PIDN:CAA38362.1; PID:gl7711

A;Note: the authors translated the codon CTT for residue 13 as Val and GTC for residue

R;Bassuener, R.; Schlesier, B.

submitted to the EMBL Data Library, December 1990

A;Reference number: S14479

A;Accession: S14479

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-146 <BAS>

A;Cross-references: EMBL:X57027; NID:gl7714; PID:gl7715

R;Altendorf, S.B.; Pearson, K.W.; Leung, F.W.; Sun, S.S.M.

Plant Mol. Biol. 8, 239-250, 1987

A;Title: Cloning and sequence analysis of a cDNA encoding a Brazil nut protein exceptic

A;Reference number: S06252

A;Accession: S06252

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-146 <ALT>

A;Experimental source: clone pHS-3

A;Note: part of this sequence, including the amino end of the large chain, was confirme

R;Bassuener, R.; Schlesier, B.

submitted to the EMBL Data Library, December 1990

A;Reference number: S21640

A;Accession: S21640

A;Status: Preliminary

A;Molecule type: DNA

A;Residues: 1-146 <BA2>

A;Cross-references: EMBL:X57028; NID:gl7716; PID:gl7717

R;Ampe, C.; Van Damme, J.; de Castro, L.A.B.; Sampaio, M.J.A.M.; Van Montagu, M.; Vande

Eur. J. Biochem. 159, 597-604, 1986

A;Title: The amino-acid sequence of the 2S sulphur-rich proteins from seeds of Brazil r

A;Reference number: A91173; MUID:87004679; PMID:3758080

A;Accession: B25802

A;Molecule type: Protein

A;Residues: 70-90, E', 92-121, 'M', 123-125, 'L', 127-142 <AMP>

C;Genetics:

A;Introns: 60/3

C;Superfamily: wheat alpha-amylase inhibitor

F;1-22/Domain: signal sequence #status predicted <SIG>
 F;23-46/Domain: propeptide #status predicted <PRO>
 F;47-69/Product: seed storage protein small chain #status predicted <SCH>
 F;70-146/Product: seed storage protein large chain #status experimental <LCH>

Query Match 79.4%; Score 27; DB 2; Length 146;
 Best Local Similarity 83.3%; Pred. No. 18;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MWMML 7
 Db 99 MWMML 104

RESULT 7

JQ1320
 high mobility group protein PF16 - malaria parasite (Plasmodium falciparum)
 C;Species: Plasmodium falciparum
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jun-2000
 C;Accession: JQ1320
 R;Guntaka, R.V.; Kandala, J.C.; Reddy, V.D.
 Biochem. Biophys. Res. Commun. 182, 412-419, 1992
 A;Title: Cloning and characterization of a highly conserved HMG-like protein (PF16) gene
 A;Reference number: JQ1320; MUID:92118046; PMID:1731798
 A;Accession: JQ1320
 A;Molecule type: DNA
 A;Residues: 1-147 <GUN>
 A;Cross-references: GB:M86518; NID:gl60325; PID:gl60326
 A;Experimental source: strain FCH/3
 C;Comment: This protein interacts with other nuclear proteins and serve as a transcript

Query Match 79.4%; Score 27; DB 2; Length 147;
 Best Local Similarity 83.3%; Pred. No. 18;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWMML 6
 Db 11 MWMML 16

RESULT 8

T04869
 transforming protein myb homolog F28A21.180 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 19-May-2000
 C;Accession: T04869
 R;Bevan, M.; Mueller, M.W.; Muendlein, A.; Felber, R.; Bancroft, I.; Mewes, H.W.; Mayer,
 submitted to the Protein Sequence Database, February 1999
 A;Reference number: Z15387
 A;Accession: T04869
 A;Molecule type: DNA
 A;Residues: 1-427 <BEV>
 A;Cross-references: EMBL:AL035526
 A;Experimental source: cultivar Columbia; BAC clone F28A21
 C;Genetics:
 A;Map position: 4
 A;Introns: 226/2; 268/3
 A;Note: F28A21.180
 C;Superfamily: unassigned myb DNA-binding repeat proteins; myb DNA-binding repeat homolo

Query Match 79.4%; Score 27; DB 2; Length 427;
 Best Local Similarity 83.3%; Pred. No. 47;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MWMML 7
 Db 145 MWMML 150

RESULT 9

A46227
 voltage-dependent Ca2+ channel alpha 1-subunit - golden hamster
 C;Species: Mesocricetus auratus (golden hamster)
 C;Date: 27-Oct-1993 #sequence_revision 18-Nov-1994 #text_change 11-Jan-2000
 C;Accession: A46227
 R;Yanay, G.C.; Wheeler, M.B.; Wei, X.; Perez-Reyes, E.; Birnbaumer, L.; Boyd III, A.E.;
 Mol. Endocrinol. 6, 2143-2152, 1992
 A;Title: Cloning of a novel alpha 1-subunit of the voltage-dependent calcium channel f
 A;Reference number: A46227; MUID:93149124; PMID:1337146
 A;Accession: A46227
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-1610 <YAN>

A;Experimental source: insulin-secreting cell line HIT-T15
 A;Note: sequence extracted from NCBI backbone (NCBI:P:123692)
 C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 79.4%; Score 27; DB 2; Length 1610;
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWMML 6
 Db 2 MWMML 7

RESULT 10

JH0422

voltage-dependent calcium channel complex alpha-1 chain - rat

C;Species: Rattus norvegicus (Norway rat)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Jan-2000
 C;Accession: JH0422; D35901; I60901
 R;Hui, A.; Ellinor, P.T.; Krizanov, O.; Wang, J.J.; Diebold, R.J.; Schwartz, A.
 Neuron 7, 35-44, 1991
 A;Title: Molecular cloning of multiple subtypes of a novel rat brain isoform of the al
 A;Reference number: JH0422; MUID:91299338; PMID:1648940
 A;Accession: JH0422
 A;Molecule type: mRNA
 A;Residues: 1-1646 <HUI>

A;Cross-references: GB:M57682; NID:g206573; PID:AAA42015.1; PID:g206574
 A;Experimental source: brain
 R;Snutch, T.P.; Leonard, J.P.; Gilbert, M.M.; Lester, H.A.; Davidson, N.
 Proc. Natl. Acad. Sci. U.S.A. 87, 3391-3395, 1990
 A;Title: Rat brain expresses a heterogeneous family of calcium channels.
 A;Reference number: A35901; MUID:90239020; PMID:1692134
 A;Accession: D35901
 A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr

A;Molecule type: mRNA
 A;Residues: 1247-1434 <SNU>
 A;Experimental source: brain
 R;Yu, A.S.L.; Hebert, S.C.; Brenner, B.M.; Lytton, J.
 Proc. Natl. Acad. Sci. U.S.A. 89, 10494-10498, 1992
 A;Title: Molecular characterization and nephron distribution of a family of transcripts

A;Reference number: A46422; MUID:93066265; PMID:1279681
 A;Accession: I60901

A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1040-1261.1305-1365 <RES>

A;Cross-references: GB:M99221; NID:g203370; PID:AAA40895.1; PID:g203371
 A;Experimental source: kidney
 C;Comment: Calcium channels are essential for many cellular functions, such as muscle c

C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain
 C;Keywords: alternative splicing; calcium binding; calcium channel; glycoprotein; phosph
 F;1463-1491/Domain: calcium binding #status predicted <EPC>
 F;154-224-328/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;464-948-1489-1584/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 79.4%; Score 27; DB 2; Length 1646;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
Db 2 M M M M M M K 7

RESULT 11
JH0564
calcium channel alpha-1D chain - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 24-Nov-1999
R:Williams, M.E.; Feldman, D.H.; McCue, A.F.; Brenner, R.; Velicelebi, G.; Ellis, S.B.;
Neuron 8, 71-84, 1992
A:Title: Structure and functional expression of alpha1, alpha2, and beta subunits of a n
A:Reference number: JH0564; MUID:92110010; PMID:1309651
A:Accession: JH0564
A:Molecule type: mRNA
A:Residues: 1-2161 <MIL>
A:Cross-references: GB:M76558
A:Experimental source: neuroblastoma, cell line IMR32
C:Comment: This protein is a subunit of the voltage-dependent calcium channel.
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain
C:Keywords: transmembrane protein
F:127-145/Domain: transmembrane #status predicted <IS1>
F:164-183/Domain: transmembrane #status predicted <IS2>
F:195-214/Domain: transmembrane #status predicted <IS3>
F:236-254/Domain: transmembrane #status predicted <IS4>
F:274-293/Domain: transmembrane #status predicted <IS5>
F:384-406/Domain: transmembrane #status predicted <IS6>
F:524-542/Domain: transmembrane #status predicted <II1>
F:558-577/Domain: transmembrane #status predicted <II2>
F:598-603/Domain: transmembrane #status predicted <II3>
F:615-633/Domain: transmembrane #status predicted <II4>
F:653-672/Domain: transmembrane #status predicted <II5>
F:728-752/Domain: transmembrane #status predicted <II6>
F:888-905/Domain: transmembrane #status predicted <S11>
F:922-941/Domain: transmembrane #status predicted <S21>
F:954-972/Domain: transmembrane #status predicted <S31>
F:980-998/Domain: transmembrane #status predicted <S41>
F:1018-1037/Domain: transmembrane #status predicted <S51>
F:1128-1152/Domain: transmembrane #status predicted <S61>
F:1206-1224/Domain: transmembrane #status predicted <S71>
F:1240-1259/Domain: transmembrane #status predicted <S81>
F:1268-1286/Domain: transmembrane #status predicted <S91>
F:1315-1333/Domain: transmembrane #status predicted <V34>
F:1353-1372/Domain: transmembrane #status predicted <V35>
F:1440-1464/Domain: transmembrane #status predicted <V36>

Query Match 79.4%; Score 27; DB 2; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
Db 3 M M M M M M K 8

RESULT 12
A38198
calcium channel alpha-1 chain, pancreatic - human
N:Alternate names: beta-cell-type calcium channel alpha-1 chain; neuroendocrine-type cal
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000
C:Accession: A38198
R:Seino, S.; Chen, L.; Seino, M.; Blondel, O.; Takeda, J.; Johnson, J.H.; Bell, G.I.
Proc. Natl. Acad. Sci. U.S.A. 89, 584-588, 1992
A:Title: Cloning of the alpha 1 subunit of a voltage-dependent calcium channel expressed
A:Reference number: A38198; MUID:92115705; PMID:1309948
A:Accession: A38198
A:Molecule type: mRNA
A:Residues: 1-2181 <SEI>

A:Cross-references: GB:M83566; NID:G179751; PIDN:AAA35629.1; PID:G179752
A:Experimental source: pancreatic beta cells
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain
C:Keywords: membrane protein; voltage-gated ion channel

Query Match 79.4%; Score 27; DB 2; Length 2181;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
Db 3 M M M M M M K 8

RESULT 13
T42742
voltage-dependent calcium channel alpha 1 chain, isoform CACM4A - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C:Accession: T42742
R:Ihara, Y.; Yamada, Y.; Fujii, Y.; Gonori, T.; Yano, H.; Yasuda, K.; Inagaki, N.; Seir
Mol. Endocrinol. 9, 121-130, 1995
A:Title: Molecular diversity and functional characterization of voltage-dependent calci
A:Reference number: Z22258; MUID:95280950; PMID:7760845
A:Accession: T42742
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2203 <IHA>
A:Cross-references: EMBL:D38101; NID:G736711; PIDN:BAA07282.1; PID:G736712
A:Experimental source: insulinoma RIN5F complementary DNA library
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain
C:Keywords: calcium channel

Query Match 79.4%; Score 27; DB 2; Length 2203;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
Db 3 M M M M M M K 8

RESULT 14
H82695
DNA-3-methyladenine glycosidase XF1326 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: H82695
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequ
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A83515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: H82695
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-216 <SIM>
A:Cross-references: GB:AE003965; GB:AE003849; NID:G9106313; PIDN:AAF84135.1; GSPDB:GN0C
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer,
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froi
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.B.; Lai
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
F.G.; Nunes, L.N.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D
Rodrigues, V.; Rosa, A.U. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawas
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve
M.; Tshakko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Contents: annotation

Thu Feb 12 09:07:34 2004

us-09-901-187c-12.rpr

C;Genetics:
A;Gene: XF1326

Query Match 76.5%; Score 26; DB 2; Length 216;
Best Local Similarity 71.4%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNNMMRL 7
|||
Db 139 MMLMRL 145

RESULT 15
H96811
protein F3F9.20 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: H96811
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H96811
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-302 <STO>
A;Cross-references: GB:AE005173; NID:G8052542; PIDN:AAF71806.1; GSPDB:GN00141
C;Genetics:
A;Gene: F3F9.20
A;Map position: 1

Query Match 76.5%; Score 26; DB 2; Length 302;
Best Local Similarity 71.4%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNNMMRL 7
|||
Db 159 MNNMMML 165

Search completed: February 11, 2004, 17:11:56
Job time : 11.3333 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 Seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-12

Perfect score: 34
Sequence: 1 MWMWRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	27	79.4	10	US-08-417-279-2	Sequence 2, Appli
2	27	79.4	10	PCT-US94-14173-2	Sequence 2, Appli
3	27	79.4	61	US-09-252-991A-29802	Sequence 29802, A
4	27	79.4	2161	US-07-745-206A-2	Sequence 2, Appli
5	27	79.4	2161	US-08-455-543A-49	Sequence 49, Appli
6	27	79.4	2161	US-08-455-543A-51	Sequence 51, Appli
7	27	79.4	2161	US-08-223-305C-49	Sequence 49, Appli
8	27	79.4	2161	US-08-223-305C-51	Sequence 51, Appli
9	27	79.4	2161	US-08-311-363-2	Sequence 2, Appli
10	26	76.5	385	US-08-891-254-3	Sequence 3, Appli
11	26	76.5	385	US-08-819-539-3	Sequence 3, Appli
12	26	76.5	385	PCT-US93-06243-2	Sequence 3, Appli
13	26	76.5	385	PCT-US96-08819-3	Sequence 3, Appli
14	26	76.5	403	US-08-200-724A-2	Sequence 2, Appli
15	26	76.5	403	US-09-030-270A-3	Sequence 3, Appli
16	26	76.5	403	US-08-851-376A-2	Sequence 2, Appli
17	26	76.5	403	US-08-984-207-3	Sequence 3, Appli
18	26	76.5	403	US-09-013-587-3	Sequence 3, Appli
19	26	76.5	403	US-09-086-118-23	Sequence 23, Appli
20	26	76.5	989	US-08-070-301-16	Sequence 16, Appli
21	25	73.5	62	US-08-488-961-6	Sequence 6, Appli
22	25	73.5	62	US-08-973-297-6	Sequence 6, Appli
23	25	73.5	62	PCT-US96-06511-6	Sequence 6, Appli
24	25	73.5	67	US-09-120-365-98	Sequence 98, Appli
25	25	73.5	67	US-09-515-039-98	Sequence 98, Appli
26	25	73.5	185	US-09-328-352-6508	Sequence 6508, Ap
27	25	73.5	284	US-09-328-352-4466	Sequence 4466, Ap

28	25	73.5	521	4	US-09-107-532A-3961	Sequence 3961, Ap
29	25	73.5	935	1	US-07-707-367-2	Sequence 2, Appli
30	25	73.5	1160	4	US-09-328-352-6457	Sequence 6457, Ap
31	24	70.6	110	4	US-09-252-991A-16573	Sequence 16573, A
32	24	70.6	190	4	US-09-328-352-4549	Sequence 4549, Ap
33	24	70.6	299	4	US-09-352-991A-22835	Sequence 22835, A
34	24	70.6	312	4	US-09-107-532A-6637	Sequence 6637, Ap
35	24	70.6	393	4	US-09-252-991A-16771	Sequence 16771, A
36	24	70.6	429	4	US-09-252-991A-16841	Sequence 16841, A
37	24	70.6	490	4	US-09-252-991A-28975	Sequence 28975, A
38	24	70.6	668	4	US-09-252-991A-22341	Sequence 22341, A
39	24	70.6	687	2	US-08-540-804-4	Sequence 4, Appli
40	24	70.6	687	2	US-08-218-265-4	Sequence 4, Appli
41	24	70.6	687	3	US-08-521-872-4	Sequence 4, Appli
42	24	70.6	687	3	US-08-590-399-4	Sequence 5555, Ap
43	23	67.6	88	4	US-09-328-352-5555	Sequence 5186, Ap
44	23	67.6	93	4	US-09-134-001C-5186	Sequence 3, Appli
45	23	67.6	189	3	US-09-105-343A-3	

ALIGNMENTS

RESULT 1
US-08-417-279-2
; Sequence 2, Application US/08417279
; Patent No. 5543498
; GENERAL INFORMATION:
; APPLICANT: Fishman, Mark C.
; APPLICANT: Igarashi, Michihiro
; TITLE OF INVENTION: Peptides to Overcome Inhibition of Nerve
; TITLE OF INVENTION: Growth
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,279
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/166,350
; FILING DATE: 14-DEC-1993
; APPLICATION NUMBER: 08/162,480
; FILING DATE: 07-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 0609.3960001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; US-08-417-279-2

Query Match 79.4%; Score 27; DB 1; Length 10;
Best Local Similarity 83.3%; Pred. No. 0.94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWMWRL 6

APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-52517
TELEPHONE: (619)238-0959
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2161 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-455-543A-49

Query Match 79.4%; Score 27; DB 1; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNNNNRR 6
| | | | |
Db 3 MNNNNRK 8

US-08-455-543A-51
Sequence 51, Application US/08455543A
Patent No. 5792846
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,543A
FILING DATE: May 31, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-52517
TELEPHONE: (619)238-0959
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 2161 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-455-543A-51

Query Match 79.4%; Score 27; DB 1; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNNNNRR 6
| | | | |
Db 3 MNNNNRK 8

US-08-455-543A-51

Qy 1 M M M M M R 6
| | | | |
Db 3 M M M M M K 8

RESULT 7

US-08-223-305C-49
; Sequence 49, Application US/08223305C
; Patent No. 5851824
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/223,305C
; FILING DATE: April 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/868,354
; FILING DATE: April 10, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US89/01408
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/176,899
; FILING DATE: 04-APR-1988
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 52516 (P519739)
; TELEPHONE: (619)238-0062
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2161 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-223-305C-49

Query Match

79.4%; Score 27; DB 2; Length 2161;

Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 M M M M M R 6
| | | | |
Db 3 M M M M M K 8

RESULT 8

US-08-223-305C-51
; Sequence 51, Application US/08223305C
; Patent No. 5851824
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/223,305C
; FILING DATE: April 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/868,354
; FILING DATE: April 10, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US89/01408
; FILING DATE: 04-APR-1989
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 52516 (P519739)
; TELEPHONE: (619)238-0062
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2161 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal

US-08-223-305C-51

Query Match 79.4%; Score 27; DB 2; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M R 6
DB 3 M M M M M K 8

RESULT 9

US-08-311-363-2
Sequence 2, Application US/08311363
Patent No. 5876958
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: Human Calcium Channel Compositions and
TITLE OF INVENTION: Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,363
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-51506
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0099
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2161 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-311-363-2

Query Match 79.4%; Score 27; DB 2; Length 2161;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M R 6
DB 3 M M M M M K 8

RESULT 10

US-08-891-254-3
Sequence 3, Application US/08891254
Patent No. 5776889
GENERAL INFORMATION:
APPLICANT: Wei, Zhong-Min

APPLICANT: Beer, Steven V.
TITLE OF INVENTION: Hypersensitive Response
TITLE OF INVENTION: Induced Resistance In Plants
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/891,254
FILING DATE: 10-JUL-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/475,775
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 14603/10050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-891-254-3

Query Match 76.5%; Score 26; DB 1; Length 385;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 M M M M M R L 7
DB 63 M M M M M M S M 69

RESULT 11

US-08-819-539-3
Sequence 3, Application US/08819539
Patent No. 5859324
GENERAL INFORMATION:
APPLICANT: Wei, Zhong-Min
APPLICANT: Beer, Steven V.
TITLE OF INVENTION: Hypersensitive Response
TITLE OF INVENTION: Induced Resistance In Plants
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/819,539
FILING DATE: 17-MAR-1997

CLASSIFICATION: 800
 PRIOR APPLICATION DATA: 08/475,775
 FILING DATE: 01-JUL-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Goldman, Michael L.
 REGISTRATION NUMBER: 30,727
 REFERENCE/DOCKET NUMBER: 14603/10050
 TELEPHONE: (716) 263-1304
 TELEFAX: (716) 263-1600
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 385 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-819-539-3

Query Match 76.5%; Score 26; DB 2; Length 385;
 Best Local Similarity 71.4%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MWMWML 7
 DB 63 MWMWMSM 69

RESULT 12
 PCT-US93-06243-2
 ; Sequence 2, Application PC/TUS9306243
 ; GENERAL INFORMATION:
 ; APPLICANT: Zhong-Min Wei, David W. Bauer, Steven V.
 ; APPLICANT: Beer, Alan Collmer, Sheng-Yang He, and Ron J. Laby
 ; TITLE OF INVENTION: Elicitor of the Hypersensitive Response in Plants
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Yahwak & Associates
 ; STREET: 25 Skytop Drive
 ; CITY: Trumbull
 ; STATE: Connecticut
 ; COUNTRY: USA
 ; ZIP: 06611
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: Microsoft Word 4.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/06243
 ; FILING DATE: 19930630
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 907,935
 ; FILING DATE: 01-JUL-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: George M. Yahwak
 ; REGISTRATION NUMBER: 26,824
 ; REFERENCE/DOCKET NUMBER: CRF D-1172
 ; TELEPHONE: (203) 268-1951
 ; TELEFAX: (203) 268-1951
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 385 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; PCT-US93-06243-2

Query Match 76.5%; Score 26; DB 5; Length 385;

PCT-US93-06243-2

Best Local Similarity 71.4%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MWMWML 7
 DB 63 MWMWMSM 69

RESULT 13
 PCT-US96-08819-3
 ; Sequence 3, Application PC/TUS9608819
 ; GENERAL INFORMATION:
 ; APPLICANT: Cornell Research Foundation, Inc.
 ; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED
 ; TITLE OF INVENTION: RESISTANCE IN PLANTS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
 ; STREET: Clinton Square, P.O. Box 1051
 ; CITY: Rochester
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 14603
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US96/08819
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/475,775
 ; FILING DATE: 07-JUN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Goldman, Michael L.
 ; REGISTRATION NUMBER: 30,727
 ; REFERENCE/DOCKET NUMBER: 19603/10051
 ; TELEPHONE: (716) 263-1304
 ; TELEFAX: (716) 263-1600
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 385 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US96-08819-3

Query Match 76.5%; Score 26; DB 5; Length 385;
 Best Local Similarity 71.4%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MWMWML 7
 DB 63 MWMWMSM 69

RESULT 14
 US-08-200-724A-2
 ; Sequence 2, Application US/08200724A
 ; Patent No. 5849868
 ; GENERAL INFORMATION:
 ; APPLICANT: Wei, Zhong-Min
 ; APPLICANT: Bauer, David W.
 ; APPLICANT: Beer, Steven V.
 ; APPLICANT: Collmer, Alan
 ; APPLICANT: He, Sheng-Yang
 ; APPLICANT: Laby, Ron J.
 ; TITLE OF INVENTION: ELICITOR OF THE HYPERSENSITIVE RESPONSE
 ; TITLE OF INVENTION: IN PLANTS

Query Match 76.5%; Score 26; DB 5; Length 385;
 Best Local Similarity 71.4%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MWMWML 7
 DB 63 MWMWMSM 69

RESULT 14
 US-08-200-724A-2
 ; Sequence 2, Application US/08200724A
 ; Patent No. 5849868
 ; GENERAL INFORMATION:
 ; APPLICANT: Wei, Zhong-Min
 ; APPLICANT: Bauer, David W.
 ; APPLICANT: Beer, Steven V.
 ; APPLICANT: Collmer, Alan
 ; APPLICANT: He, Sheng-Yang
 ; APPLICANT: Laby, Ron J.
 ; TITLE OF INVENTION: ELICITOR OF THE HYPERSENSITIVE RESPONSE
 ; TITLE OF INVENTION: IN PLANTS

NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/200,724A
APPLICATION NUMBER: US/08/200,724A
FILING DATE: 23-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/10030
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-200-724A-2

Query Match 76.5%; Score 26; DB 2; Length 403;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 M M M M M R L 7
| | | | | :
D b 63 M M M M M S M 69

RESULT 15
US-09-030-270A-3
Sequence 3, Application US/09030270A
Patent No. 5977060
GENERAL INFORMATION:
APPLICANT: Zitter, Thomas A.
APPLICANT: Wei, Zhong-Min
TITLE OF INVENTION: INSECT CONTROL WITH A
TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,270A
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/039,226
FILING DATE: 28-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/1521
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-030-270A-3

Query Match 76.5%; Score 26; DB 2; Length 403;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 M M M M M R L 7
| | | | | :
D b 63 M M M M M S M 69

Search completed: February 11, 2004, 17:13:40
Job time : 11.4167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.5667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-12

Perfect score: 34
Sequence: 1 MAMMRL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	7	10	US-09-901-187B-12
2	32	94.1	103	9	US-09-864-761-44210
3	29	85.3	42	9	US-09-864-761-45051
4	28	82.4	32	9	US-09-864-761-34021
5	28	82.4	223	12	US-10-108-2608-4500
6	28	82.4	255	12	US-10-017-161-2180
7	28	82.4	255	12	US-10-292-798-1826
8	27	79.4	18	9	US-09-864-761-35316
9	27	79.4	18	9	US-09-864-761-38185
10	27	79.4	33	9	US-09-864-761-45071
11	27	79.4	111	9	US-09-864-761-43445
12	27	79.4	305	10	US-09-866-055-35
13	27	79.4	305	11	US-09-779-679-14
14	27	79.4	305	11	US-09-779-679-16
15	27	79.4	305	11	US-09-779-679-18

16	79.4	305	11	US-09-779-679-22	Sequence 22, Appl
17	79.4	305	11	US-09-804-291-35	Sequence 35, Appl
18	79.4	305	12	US-10-387-629-42	Sequence 42, Appl
19	79.4	427	12	US-10-374-780A-2722	Sequence 2722, Ap
20	79.4	2161	12	US-10-375-283-2	Sequence 2, Appl
21	79.4	2166	14	US-10-029-413A-4	Sequence 4, Appl
22	79.4	2181	12	US-10-411-010-29	Sequence 29, Appl
23	79.4	2181	14	US-10-029-413A-18	Sequence 18, Appl
24	79.4	2181	14	US-10-029-413A-20	Sequence 20, Appl
25	79.4	2181	15	US-10-205-823-50	Sequence 50, Appl
26	76.5	18	9	US-09-864-761-35295	Sequence 35295, A
27	76.5	20	9	US-09-864-761-33376	Sequence 33376, A
28	76.5	22	9	US-09-864-761-40486	Sequence 40486, A
29	76.5	25	9	US-09-864-761-33850	Sequence 33850, A
30	76.5	26	9	US-09-864-761-42400	Sequence 42400, A
31	76.5	27	9	US-09-864-761-40516	Sequence 40516, A
32	76.5	27	9	US-09-864-761-41279	Sequence 41279, A
33	76.5	33	9	US-09-864-761-48633	Sequence 48633, A
34	76.5	50	9	US-09-864-761-44258	Sequence 44258, A
35	76.5	60	12	US-10-029-386-31671	Sequence 31671, A
36	76.5	69	9	US-09-864-761-42285	Sequence 42285, A
37	76.5	403	9	US-09-866-118-23	Sequence 23, Appl
38	76.5	403	9	US-09-835-684-3	Sequence 3, Appl
39	76.5	403	9	US-09-880-371-3	Sequence 3, Appl
40	76.5	403	9	US-09-879-248-3	Sequence 3, Appl
41	76.5	403	9	US-09-770-693-3	Sequence 3, Appl
42	76.5	403	10	US-09-766-348-3	Sequence 3, Appl
43	76.5	403	12	US-10-387-806-23	Sequence 23, Appl
44	76.5	403	12	US-10-441-736-3	Sequence 3, Appl
45	76.5	403	15	US-10-034-158-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-901-187B-12
; Sequence 12, Application US/09901187B
; Patent No. US20020151464A1
; GENERAL INFORMATION:
; APPLICANT: Panacea Pharmaceuticals, Inc.
; APPLICANT: Wolozin, Benjamin
; APPLICANT: Ostretova-Golts, Natalie
; APPLICANT: Lebowitz, Micheal S
; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatment
; FILE REFERENCE: PAN01/002US
; CURRENT APPLICATION NUMBER: US/09/901.187B
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/217,319
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/279,199
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 12
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-187B-12

Query Match 100.0%; Score 34; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMMRL 7

Db 1 MAMMRL 7

RESULT 2

US-09-864-761-44210
; Sequence 44210, Application US/09864761

RESULT 4
US-09-864-761-34021
; Sequence 34021, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34021
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005100.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.74
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.81
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.6
US-09-864-761-34021

Query Match 82.4%; Score 28; DB 9; Length 32;
Best Local Similarity 85.7%; Pred. No. 9.2;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 M M M M M R L 7
| | | | |
DB 3 M M M M M M L 9
| | | | |
RESULT 5
US-10-108-260A-4500
; Sequence 4500, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4500
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4500

Query Match 82.4%; Score 28; DB 12; Length 223;
Best Local Similarity 85.7%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 M M M M M R L 7
| | | | |
DB 1 M M M M M M L 7
| | | | |

RESULT 6
US-10-017-161-2180
; Sequence 2180, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2180
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-2180

Query Match 82.4%; Score 28; DB 12; Length 255;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M R L 7
| | | | |
DB 111 M M M M M V R M 117
| | | | |

RESULT 7
US-10-292-798-1826
; Sequence 1826, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI

;; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS

;; FILE REFERENCE: 084335/166
;; CURRENT APPLICATION NUMBER: US/10/292,798
;; CURRENT FILING DATE: 2002-11-13
;; PRIOR APPLICATION NUMBER: 10/017,161
;; PRIOR FILING DATE: 2001-12-18
;; PRIOR APPLICATION NUMBER: JP 2001-246789
;; PRIOR FILING DATE: 2001-06-18
;; NUMBER OF SEQ ID NOS: 2070
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO: 1826
;; LENGTH: 255
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-292-798-1826

Query Match 82.4%; Score 28; DB 12; Length 255;

Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMMMMRL 7
Db 111 MMMVM 117

RESULT 8

US-09-864-761-35316
;; Sequence 35316, Application US/09864761
;; Patent No. US20020048763A1

GENERAL INFORMATION:

;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aecmca-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23

;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30

;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
;; SEQ ID NO: 35316
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL035704.7
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
US-09-864-761-35316

Query Match 79.4%; Score 27; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 8.8;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMMMMR 6
Db 13 MMMMMK 18

RESULT 9

US-09-864-761-38185
;; Sequence 38185, Application US/09864761
;; Patent No. US20020048763A1

GENERAL INFORMATION:

;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aecmca-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23

;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670

;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 38185
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL035704.9
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.1
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN HELL100, SIGNAL = 2.8
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.1
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
;;
US-09-864-761-38185

Query Match 79.4%; Score 27; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 8.8;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMR 6
 |||||
DB 13 MMMMMK 18

RESULT 10
US-09-864-761-45071
;; Sequence 45071, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aecmca-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663

;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 45071
;; LENGTH: 33
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC025644.2
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.84
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.59
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.76
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.57
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.57
;;
US-09-864-761-45071

Query Match 79.4%; Score 27; DB 9; Length 33;
Best Local Similarity 83.3%; Pred. No. 16;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMR 6
 |||||
DB 10 MMMMMK 15

RESULT 11
US-09-864-761-43445
;; Sequence 43445, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aecmca-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 43445
;; LENGTH: 111
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006097.1
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.89
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.62
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.6
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.81
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
;; OTHER INFORMATION: SWISSPROT HIT: Q10699, EVALUATE 5.90e-01
US-09-864-761-43445

Query Match 79.4%; Score 27; DB 9; Length 111;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMR 6
Db 1 MMMMMK 6

RESULT 12

US-09-886-055-35
;; Sequence 35, Application US/09886055
;; Patent No. US20020132273A1
;; GENERAL INFORMATION:
;; APPLICANT: STRYER, LUBERT
;; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
;; BIOSENSORS OF CHEMICAL SENSANTS
;; FILE REFERENCE: 078003-0277150
;; CURRENT APPLICATION NUMBER: US/09/886,055
;; PRIOR FILING DATE: 2001-06-22
;; PRIOR APPLICATION NUMBER: 60/213,812
;; PRIOR FILING DATE: 2000-06-22
;; NUMBER OF SEQ ID NOS: 522
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 35
;; LENGTH: 305
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-886-055-35

Query Match 79.4%; Score 27; DB 10; Length 305;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMR 7
Db 45 MIMLMR 51

RESULT 13

US-09-779-679-14

;; Sequence 14, Application US/09779679
;; Publication No. US20030082757A1
;; GENERAL INFORMATION:
;; APPLICANT: Taupier, Raymond J
;; APPLICANT: Burgess, Catherine E
;; APPLICANT: Padigaru, Muralidhara
;; APPLICANT: Tchernev, Velizar T
;; APPLICANT: Mishra, Vishnu S
;; APPLICANT: Casman, Stacie
;; APPLICANT: Ballinger, Robert
;; APPLICANT: Vernet, Corine A
;; APPLICANT: Li, Li
;; APPLICANT: Spytek, Kimberly A
;; APPLICANT: Andrew, David P
;; APPLICANT: Mezes, Peter S
;; TITLE OF INVENTION: No. US20030082757A1 Proteins and Nucleic Acids Encoding the Sa
;; FILE REFERENCE: 15966-661
;; CURRENT APPLICATION NUMBER: US/09/779,679
;; PRIOR FILING DATE: 2001-02-08
;; PRIOR APPLICATION NUMBER: USN 60/181045
;; PRIOR FILING DATE: 2000-02-08
;; PRIOR APPLICATION NUMBER: USN 60/183191
;; PRIOR FILING DATE: 2000-02-17
;; PRIOR APPLICATION NUMBER: USN 60/180,929
;; PRIOR FILING DATE: 2000-02-08
;; PRIOR APPLICATION NUMBER: USN 60/219758
;; PRIOR FILING DATE: 2000-07-20
;; PRIOR APPLICATION NUMBER: USN 60/181339
;; PRIOR FILING DATE: 2000-02-09
;; PRIOR APPLICATION NUMBER: USN 60/181344
;; PRIOR FILING DATE: 2000-02-09
;; PRIOR APPLICATION NUMBER: USN 60/221341
;; PRIOR FILING DATE: 2000-07-26
;; PRIOR APPLICATION NUMBER: USN 60/181392
;; PRIOR FILING DATE: 2000-02-09
;; PRIOR APPLICATION NUMBER: USN 60/219585
;; PRIOR FILING DATE: 2000-07-20
;; PRIOR APPLICATION NUMBER: USN 60/181157
;; PRIOR FILING DATE: 2000-02-09
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 14
;; LENGTH: 305
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-779-679-14

Query Match 79.4%; Score 27; DB 11; Length 305;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMR 7
Db 45 MIMLMR 51

RESULT 14

US-09-779-679-16
;; Sequence 16, Application US/09779679
;; Publication No. US20030082757A1
;; GENERAL INFORMATION:
;; APPLICANT: Taupier, Raymond J
;; APPLICANT: Burgess, Catherine E
;; APPLICANT: Padigaru, Muralidhara
;; APPLICANT: Tchernev, Velizar T
;; APPLICANT: Mishra, Vishnu S
;; APPLICANT: Casman, Stacie
;; APPLICANT: Ballinger, Robert
;; APPLICANT: Vernet, Corine A
;; APPLICANT: Li, Li
;; APPLICANT: Spytek, Kimberly A
;; APPLICANT: Andrew, David P
;; APPLICANT: Mezes, Peter S

; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sam
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-679-16

Query Match 79.4%; Score 27; DB 11; Length 305;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMRL 7
DB 45 MIMLMRL 51

RESULT 15
US-09-779-679-18
; Sequence 18, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sam
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344

; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 305.
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-679-18

Query Match 79.4%; Score 27; DB 11; Length 305;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMMMMRL 7
DB 45 MIMLMRL 51

Search completed: February 11, 2004, 17:54:14
Job time : 25.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 seconds
(without alignments)
34.452 Million cell updates/sec

Title: US-09-901-187C-12

Perfect score: 34

Sequence: 1 MMMWRL 7

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

A_Geneseq_15Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	7	23	AAE14557
2	32	94.1	103	22	ABG55058
3	32	94.1	103	22	ABH3925
4	32	94.1	103	22	AAW60670
5	32	94.1	103	22	AAW73342
6	32	94.1	103	22	AAW33544
7	32	94.1	103	23	ABG43195
8	30	88.2	167	21	ABG24998
9	30	88.2	193	21	AAG51419

10	30	88.2	194	21	AAG51418	Arabidopsis thalia
11	30	88.2	212	21	AAQ24997	Arabidopsis thalia
12	30	88.2	212	21	AAQ51694	Arabidopsis thalia
13	30	88.2	214	21	AAQ51694	Arabidopsis thalia
14	30	88.2	233	21	AAQ51436	Arabidopsis thalia
15	30	88.2	235	21	AAQ51693	Arabidopsis thalia
16	30	88.2	249	21	AAQ24996	Arabidopsis thalia
17	30	88.2	249	21	AAQ51692	Arabidopsis thalia
18	30	88.2	314	21	AAQ51417	Arabidopsis thalia
19	30	88.2	334	21	AAQ23344	Arabidopsis thalia
20	30	88.2	339	21	AAQ23343	Arabidopsis thalia
21	30	88.2	344	22	AAQ40351	Human polyprotein
22	30	88.2	352	22	ABG22569	Novel human diaph
23	30	88.2	353	21	AAQ51435	Arabidopsis thalia
24	30	88.2	358	21	AAQ51434	Arabidopsis thalia
25	29	85.3	42	22	ABG57022	Human liver peptid
26	29	85.3	42	22	ABG57022	Peptide #9086 enco
27	29	85.3	42	22	ABG62451	Human brain expres
28	29	85.3	42	22	AAW75260	Human bone marrow
29	29	85.3	42	22	AAW35372	Peptide #9409 enco
30	28	82.4	32	22	ABG48110	Human liver: peptid
31	28	82.4	32	22	ABG28087	Human peptide #738
32	28	82.4	32	22	ABH33261	Peptide #767 enco
33	28	82.4	32	22	ABH18723	Protein #722 enco
34	28	82.4	32	22	AAW54053	Human brain expres
35	28	82.4	32	22	AAW14314	Peptide #748 enco
36	28	82.4	32	22	AAW26724	Peptide #761 enco
37	28	82.4	32	22	AAW02043	Peptide #725 enco
38	28	82.4	32	23	ABG36093	Human peptide enco
39	28	82.4	35	24	AAE12807	Human CEA exon #10
40	28	82.4	107	21	AAQ56156	Arabidopsis thalia
41	28	82.4	120	21	AAQ56155	Arabidopsis thalia
42	28	82.4	129	22	ABG68982	Drosophila melanog
43	28	82.4	1033	24	AAE32797	Human CEA protein
44	27	79.4	10	16	AAE3805	GAP-43 protein der
45	27	79.4	18	22	ABG49431	Human liver peptid

ALIGNMENTS

RESULT 1

AAE14557
ID AAE14557 standard; peptide; 7 AA.

XX AAE14557;

AC AAE14557;

DT 17-MAY-2002 (first entry)

XX Human alpha-synuclein aggregation inhibitor #12.

XX Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;

KW Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;

KW multiple system atrophy; Hallervorden-Spatz disease; human.

XX Homo sapiens.

OS Homo sapiens.

PN WO290204482-A1.

XX WO290204482-A1.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21379.

XX 07-JUL-2000; 2000US-217319P.

XX 28-MAR-2001; 2001US-279199P.

XX (PANA-) PANACEA PHARM INC.

XX Wolozin B, Ostretova-Golts N, Lebowitz MS;

XX WPI; 2002-179695/23.

PT Determination of an agent capable of inhibiting aggregation of alpha

PT synuclein useful for treating a neurodegenerative disease involves
 PT determining aggregation of alpha synuclein in the presence of exogenous
 PT iron or copper
 XX Claim 40; Page 37; 52pp; English.
 XX The invention relates to screening of inhibitors of alpha-synuclein
 CC aggregation in the presence of exogenous iron or copper. The inhibitors
 CC are magnesium and alpha-synuclein binding peptides, which are
 CC useful for treating neurodegenerative disease that involves
 CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
 CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
 CC atrophy and Hallervorden-Spatz disease. The present sequence is a
 CC peptide that binds to the NAC (non-amyloid-beta protein component)
 CC portion of human alpha-synuclein and inhibits its aggregation.
 XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 34; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 M M M M M R L 7
 DB 1 M M M M M R L 7
 RESULT 2
 ABG55058
 ID ABG55058 standard; Peptide; 103 AA.
 AC ABG55058;
 XX
 XX 25-FEB-2003 (first entry)
 DT Human liver peptide, SEQ ID No 33706.
 DE
 XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX
 XX WO200157273-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00664.
 P7
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488898/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver
 PT
 XX Claim 27; SEQ ID No 33706; 658pp; English.
 XX The invention relates to a single exon nucleic acid probe (SENp) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (I) may be used for predicting, measuring and displaying gene

CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent
 CC human liver single exon encoded peptides of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 103 AA;
 Query Match 94.1%; Score 32; DB 22; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 M M M M M R L 7
 DB 16 M M M M M R M 22
 RESULT 3
 ABB39925
 ID ABB39925 standard; Peptide; 103 AA.
 XX
 AC ABB39925;
 XX
 XX 04-FEB-2002 (first entry)
 DT Peptide #7431 encoded by human foetal liver single exon probe.
 DE
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KW Homo sapiens.
 XX
 XX WO200157277-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00669.
 P7
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483447/52.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human foetal liver
 PT
 XX Claim 27; SEQ ID No 32560; 639pp + sequence listing; English.
 XX The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 103 AA;
 Query Match 94.1%; Score 32; DB 22; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.7;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M L 7
 Db 16 M M M M M M R M 22

RESULT 4
 AAM60670
 ID AAM60670 standard; Protein; 103 AA.
 AC
 XX
 XX
 AC AAM60670;
 DT 05-NOV-2001 (first entry)
 XX
 XX Human brain expressed single exon probe encoded protein SEQ ID NO: 32775.
 DE Human; brain expressed exon; gene expression analysis; probe;
 XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX
 XX Homo sapiens.
 OS
 XX WO200157275-A2.
 PN
 XX
 XX 09-AUG-2001.
 PD
 XX
 XX 30-JAN-2001; 2001WO-US00667.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483446/52.
 XX
 XX Single exon nucleic acid probes for analyzing gene expression in human
 DR brains -
 XX
 XX Example 4; SEQ ID NO: 32775; 650pp + Sequence Listing; English.
 PS The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 XX
 XX Sequence 103 AA;
 SQ
 Query Match 94.1%; Score 32; DB 22; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M L 7
 Db 16 M M M M M M R M 22

RESULT 5
 AAM73342
 ID AAM73342 standard; Protein; 103 AA.
 AC
 XX
 XX AAM73342;
 XX

DT 06-NOV-2001 (first entry)
 XX Human bone marrow expressed probe encoded protein SEQ ID NO: 33648.
 DE Human; bone marrow expressed exon; gene expression analysis; probe;
 XX microarray; cancer; leukaemia; lymphoma; myeloma.
 KW
 XX Homo sapiens.
 OS
 XX WO200157276-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00668.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488900/53.
 XX
 XX Human genome-derived single exon nucleic acid probes useful for
 DR analyzing gene expression in human bone marrow -
 XX
 XX Example 4; SEQ ID NO: 33648; 658pp + Sequence Listing; English.
 PS The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 XX
 XX Sequence 103 AA;
 SQ
 Query Match 94.1%; Score 32; DB 22; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M L 7
 Db 16 M M M M M M R M 22

RESULT 6
 AAM33544
 ID AAM33544 standard; Protein; 103 AA.
 AC
 XX
 XX AAM33544;
 XX
 XX 17-OCT-2001 (first entry)
 DT Peptide #7581 encoded by probe for measuring placental gene expression.
 DE Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 XX
 XX Homo sapiens.
 OS
 XX WO200157272-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00663.
 PF
 XX

PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234487.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-48897/53.
 DR Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human placenta -
 PT
 XX Claim 27; SEQ ID No 33813; 654pp; English.
 XX
 XX The present invention relates to single exon nucleic acid probes (SENP;
 CC see AA1315-AA157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 XX
 XX Sequence 103 AA;
 SQ
 Query Match 94.1%; Score 32; DB 22; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MWMWML 7
 DB 16 MWMWML 22
 RESULT 7
 ABG43195
 ID ABG43195 standard; Peptide; 103 AA.
 AC ABG43195;
 XX
 XX 19-AUG-2002 (first entry)
 DT
 XX Human peptide encoded by genome-derived single exon probe SEQ ID 32860.
 DE
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW Primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2002-114183/15.
 DR Spatially-addressable set of single exon nucleic acid probes, used to
 XX measure gene expression in human lung samples -
 PT
 XX Claim 27; SEQ ID No 32860; 634pp; English.
 XX
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridization of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridization to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 103 AA;
 SQ
 Query Match 94.1%; Score 32; DB 23; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.7;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MWMWML 7
 DB 16 MWMWML 22
 RESULT 8
 AAG24998
 ID AAG24998 standard; Protein; 167 AA.
 XX
 AC AAG24998;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 28884.
 XX

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0128845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 06-MAY-1999; 99US-0132485.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 14-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 14-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134768.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135153.

XX 24-MAY-1999; 99US-0135629.

XX 25-MAY-1999; 99US-0136021.

XX 27-MAY-1999; 99US-0136392.

XX 28-MAY-1999; 99US-0136782.

XX 01-JUN-1999; 99US-0137222.

XX 04-JUN-1999; 99US-0137528.

XX 07-JUN-1999; 99US-0137724.

XX 08-JUN-1999; 99US-0138094.

XX 10-JUN-1999; 99US-0138540.

XX 14-JUN-1999; 99US-0138847.

XX 16-JUN-1999; 99US-0139119.

XX 16-JUN-1999; 99US-0139452.

XX 17-JUN-1999; 99US-0139453.

XX 17-JUN-1999; 99US-0139492.

XX 18-JUN-1999; 99US-0139494.

XX 18-JUN-1999; 99US-0139454.

XX 18-JUN-1999; 99US-0139455.

XX 18-JUN-1999; 99US-0139456.

XX 18-JUN-1999; 99US-0139457.

XX 18-JUN-1999; 99US-0139458.

XX 18-JUN-1999; 99US-0139459.

XX 18-JUN-1999; 99US-0139460.

XX 18-JUN-1999; 99US-0139461.

XX 18-JUN-1999; 99US-0139462.

XX 18-JUN-1999; 99US-0139463.

XX 18-JUN-1999; 99US-0139750.

XX 21-JUN-1999; 99US-0139763.

XX 22-JUN-1999; 99US-0139817.

XX 23-JUN-1999; 99US-0139899.

XX 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.

PR 27-AUG-1999; 99US-0151066.

PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.

```

PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161362.
PR 28-OCT-1999; 99US-0161392.
PR 28-OCT-1999; 99US-0161393.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 167;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNNMMNR 6
Db 92 MNNMMNR 97

RESULT 9
AAG51419
ID AAG51419 standard; Protein; 193 AA.
XX AC AAG51419;
XX AC AAG51419;
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 65257.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65257.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX Arabidopsis thaliana.
PN EPI033405-A2.

XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0128845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139859.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.

```


PR	22-SEP-1999;	99UTS-01551139
PR	23-SEP-1999;	99UTS-01551486
PR	24-SEP-1999;	99UTS-01556559
PR	25-SEP-1999;	99UTS-01556559
PR	26-SEP-1999;	99UTS-01564548
PR	27-SEP-1999;	99UTS-01565596
PR	04-OCT-1999;	99UTS-01571117
PR	05-OCT-1999;	99UTS-0157753
PR	06-OCT-1999;	99UTS-0157865
PR	07-OCT-1999;	99UTS-0158029
PR	08-OCT-1999;	99UTS-0158232
PR	12-OCT-1999;	99UTS-0158369
PR	13-OCT-1999;	99UTS-0159293
PR	13-OCT-1999;	99UTS-0159294
PR	13-OCT-1999;	99UTS-0159295
PR	14-OCT-1999;	99UTS-0159329
PR	14-OCT-1999;	99UTS-0159330
PR	14-OCT-1999;	99UTS-0159231
PR	14-OCT-1999;	99UTS-0159637
PR	14-OCT-1999;	99UTS-0159638
PR	18-OCT-1999;	99UTS-01595584
PR	21-OCT-1999;	99UTS-0160741
PR	21-OCT-1999;	99UTS-0160767
PR	21-OCT-1999;	99UTS-0160768
PR	21-OCT-1999;	99UTS-0160770
PR	21-OCT-1999;	99UTS-0160814
PR	21-OCT-1999;	99UTS-0160815
PR	22-OCT-1999;	99UTS-0160980
PR	22-OCT-1999;	99UTS-0160981
PR	22-OCT-1999;	99UTS-0160989
PR	25-OCT-1999;	99UTS-0161404
PR	25-OCT-1999;	99UTS-0161405
PR	25-OCT-1999;	99UTS-0161406
PR	26-OCT-1999;	99UTS-0161359
PR	26-OCT-1999;	99UTS-0161360
PR	26-OCT-1999;	99UTS-0161361
PR	28-OCT-1999;	99UTS-0161920
PR	28-OCT-1999;	99UTS-0161992
PR	28-OCT-1999;	99UTS-0161993
PR	29-OCT-1999;	99UTS-0162142

Qy 1 M M M M R 6
pb 1 M M M M R 6

AAG51418
ID AAG51418 standard: Protein: 194 AA.

18-OCT-2000 (first entry)

[illegible]

XX protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

33 XX

PN EP1033405-A2

☐ 1
☐ 2
☐ 3
☐ 4
☐ 5
☐ 6
☐ 7
☐ 8
☐ 9
☐ 10
☐ 11
☐ 12
☐ 13
☐ 14
☐ 15
☐ 16
☐ 17
☐ 18
☐ 19
☐ 20
☐ 21
☐ 22
☐ 23
☐ 24
☐ 25
☐ 26
☐ 27
☐ 28
☐ 29
☐ 30
☐ 31
☐ 32
☐ 33
☐ 34
☐ 35
☐ 36
☐ 37
☐ 38
☐ 39
☐ 40
☐ 41
☐ 42
☐ 43
☐ 44
☐ 45
☐ 46
☐ 47
☐ 48
☐ 49
☐ 50
☐ 51
☐ 52
☐ 53
☐ 54
☐ 55
☐ 56
☐ 57
☐ 58
☐ 59
☐ 60
☐ 61
☐ 62
☐ 63
☐ 64
☐ 65
☐ 66
☐ 67
☐ 68
☐ 69
☐ 70
☐ 71
☐ 72
☐ 73
☐ 74
☐ 75
☐ 76
☐ 77
☐ 78
☐ 79
☐ 80
☐ 81
☐ 82
☐ 83
☐ 84
☐ 85
☐ 86
☐ 87
☐ 88
☐ 89
☐ 90
☐ 91
☐ 92
☐ 93
☐ 94
☐ 95
☐ 96
☐ 97
☐ 98
☐ 99
☐ 100
☐ 101
☐ 102
☐ 103
☐ 104
☐ 105
☐ 106
☐ 107
☐ 108
☐ 109
☐ 110
☐ 111
☐ 112
☐ 113
☐ 114
☐ 115
☐ 116
☐ 117
☐ 118
☐ 119
☐ 120
☐ 121
☐ 122
☐ 123
☐ 124
☐ 125
☐ 126
☐ 127
☐ 128
☐ 129
☐ 130
☐ 131
☐ 132
☐ 133
☐ 134
☐ 135
☐ 136
☐ 137
☐ 138
☐ 139
☐ 140
☐ 141
☐ 142
☐ 143
☐ 144
☐ 145
☐ 146
☐ 147
☐ 148
☐ 149
☐ 150
☐ 151
☐ 152
☐ 153
☐ 154
☐ 155
☐ 156
☐ 157
☐ 158
☐ 159
☐ 160
☐ 161
☐ 162
☐ 163
☐ 164
☐ 165
☐ 166
☐ 167
☐ 168
☐ 169
☐ 170
☐ 171
☐ 172
☐ 173
☐ 174
☐ 175
☐ 176
☐ 177
☐ 178
☐ 179
☐ 180
☐ 181
☐ 182
☐ 183
☐ 184
☐ 185
☐ 186
☐ 187
☐ 188
☐ 189
☐ 190
☐ 191
☐ 192
☐ 193
☐ 194
☐ 195
☐ 196
☐ 197
☐ 198
☐ 199
☐ 200
☐ 201
☐ 202
☐ 203
☐ 204
☐ 205
☐ 206
☐ 207
☐ 208
☐ 209
☐ 210
☐ 211
☐ 212
☐ 213
☐ 214
☐ 215
☐ 216
☐ 217
☐ 218
☐ 219
☐ 220
☐ 221
☐ 222
☐ 223
☐ 224
☐ 225
☐ 226
☐ 227
☐ 228
☐ 229
☐ 230
☐ 231
☐ 232
☐ 233
☐ 234
☐ 235
☐ 236
☐ 237
☐ 238
☐ 239
☐ 240
☐ 241
☐ 242
☐ 243
☐ 244
☐ 245
☐ 246
☐ 247
☐ 248
☐ 249
☐ 250
☐ 251
☐ 252
☐ 253
☐ 254
☐ 255
☐ 256
☐ 257
☐ 258
☐ 259
☐ 260
☐ 261
☐ 262
☐ 263
☐ 264
☐ 265
☐ 266
☐ 267
☐ 268
☐ 269
☐ 270
☐ 271
☐ 272
☐ 273
☐ 274
☐ 275
☐ 276
☐ 277
☐ 278
☐ 279
☐ 280
☐ 281
☐ 282
☐ 283
☐ 284
☐ 285
☐ 286
☐ 287
☐ 288
☐ 289
☐ 290
☐ 291
☐ 292
☐ 293
☐ 294
☐ 295
☐ 296
☐ 297
☐ 298
☐ 299
☐ 300
☐ 301
☐ 302
☐ 303
☐ 304
☐ 305
☐ 306
☐ 307
☐ 308
☐ 309
☐ 310
☐ 311
☐ 312
☐ 313
☐ 314
☐ 315
☐ 316
☐ 317
☐ 318
☐ 319
☐ 320
☐ 321
☐ 322
☐ 323
☐ 324
☐ 325
☐ 326
☐ 327
☐ 328
☐ 329
☐ 330
☐ 331
☐ 332
☐ 333
☐ 334
☐ 335
☐ 336
☐ 337
☐ 338
☐ 339
☐ 340
☐ 341
☐ 342
☐ 343
☐ 344
☐ 345
☐ 346
☐ 347
☐ 348
☐ 349
☐ 350

PD
XX
06-SEP-2000

25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.
PR 25-FEB-1999; 99US-0123180.
PR 25-FEB-1999; 99US-0123180.

PR
05-MAR-1993;
9905-0123180

PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 194;

Best Local Similarity 100.0%; Pred. No. 9.7;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M M R 6

Db 2 M M M M M M R 7

RESULT 11

AAG24997
 ID AAG24997 standard; Protein; 212 AA.

XX AC

XX AAG24997;

XX DT

XX 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 28883.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD

XX 06-SEP-2000.

XX PF

XX 25-FEB-2000; 2000EP-0301439.

XX PR

XX 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 03-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 23-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136022.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140981.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144332.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145132.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157153.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.

PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 14-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 23-OCT-1999; 99US-0161404.
 PR 23-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 212;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M R 6
 DB 137 M M M M M R 142

RESULT 12

AAG51694
 ID AAG51694 standard; Protein; 212 AA.

XX
 AC AAG51694;

XX
 DT 18-OCT-2000 (first entry)

XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 65635.

XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 hybridisation assay; genetic mapping; gene expression control; promoter;
 termination sequence.

XX
 OS Arabidopsis thaliana.

XX
 PN EP1033405-A2.

XX
 PD 06-SEP-2000.

XX
 PF 25-FEB-2000; 2000EP-0301439.

XX
 PR 25-FEB-1999; 99US-0121825.

PR
 PR 05-MAR-1999; 99US-0123180.

PR
 PR 09-MAR-1999; 99US-0123548.

PR
 PR 23-MAR-1999; 99US-0125788.

PR
 PR 25-MAR-1999; 99US-0126264.

PR
 PR 29-MAR-1999; 99US-0126785.

PR
 PR 01-APR-1999; 99US-0127462.

PR
 PR 06-APR-1999; 99US-0128234.

PR
 PR 18-APR-1999; 99US-0128714.

PR
 PR 19-APR-1999; 99US-0130077.

PR
 PR 21-APR-1999; 99US-0130449.

PR
 PR 23-APR-1999; 99US-0130510.

PR
 PR 23-APR-1999; 99US-0130891.

PR
 PR 28-APR-1999; 99US-0131449.

PR
 PR 30-APR-1999; 99US-0132048.

PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 25-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.
 PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 212;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 M M M M M M R 6

Db 137 M M M M M R 142

RESULT 13

AAG233345

ID AAG233345 standard; Protein; 214 AA.

XX AAG233345;

XX AC AAG233345;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 26618.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 23-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 23-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 04-MAY-1999; 99US-0132407.

XX PR 05-MAY-1999; 99US-0132484.

XX PR 06-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 07-MAY-1999; 99US-0132487.

XX PR 11-MAY-1999; 99US-0132863.

XX PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135253.
 PR 21-MAY-1999; 99US-0135253.
 PR 24-MAY-1999; 99US-0136021.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.

```

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147433.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160915.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      88.2%; Score 30; DB 21; Length 214;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 M M M M M M R 6
Db 2 M M M M M M R 7

RESULT 14
AAGS1436
ID AAGS1436 standard; Protein; 233 AA.
XX AC AAGS1436;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 65281.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 23-MAR-1999; 99US-0126264.
XX 23-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134319.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.

```

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 30-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 233;
Best Local Similarity 100.0%; Pred. No. 12; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0;

Qy 1 MMTMMR 6
 |||||
Db 2 MMTMMR 7

RESULT 15
AAG51693
ID AAG51693 standard; Protein; 235 AA.
XX AC AAG51693;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 65634.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0143220.
PR 12-JUL-1999; 99US-0143277.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 18-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147535.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160982.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161923.
 PR 29-OCT-1999; 99US-0162142.

Query Match 88.2%; Score 30; DB 21; Length 235;
 Best Local Similarity 100.0%; Pred. No. 12;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 QY 1 M M M M M M M R 6
 DB 160 M M M M M M R 165

Search completed: February 11, 2004, 17:03:05
 Job time : 33.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 Seconds
(without alignments)
34.452 Million cell updates/sec

Title: US-09-901-187C-5
Perfect score: 38
Sequence: 1 THRLPSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	7	20	B. cereus T spore
2	38	100.0	7	23	Human alpha-synuclein
3	38	100.0	1608	17	Borna disease virus
4	38	100.0	1711	17	Borna disease virus
5	35	92.1	163	22	Propionibacterium
6	34	89.5	356	22	Novel human diageno
7	33	86.8	9	16	Cyclin B/p33 (cdc2)
8	33	86.8	9	21	pp60-c-src peptide
9	33	86.8	18	22	Human liver peptid

10	33	86.8	18	22	ABB28073	Human peptide #724
11	33	86.8	18	22	ABB33247	Peptide #753 encod
12	33	86.8	18	22	ABBI8710	Protein #703 encod
13	33	86.8	18	22	AA54040	Human brain expro
14	33	86.8	18	22	AA56430	Human bone marrow
15	33	86.8	18	22	AA514300	Peptide #734 encod
16	33	86.8	18	22	AA526710	Peptide #747 encod
17	33	86.8	18	22	AA526710	Peptide #711 encod
18	33	86.8	18	22	AA526710	Human peptide enco
19	33	86.8	18	22	AA526710	Human cardiovascular
20	33	86.8	18	22	AA526710	Propionibacterium
21	33	86.8	18	22	AA526710	Human secreted pro
22	33	86.8	18	22	AA526710	Human EST encod
23	33	86.8	18	22	AA526710	Propionibacterium
24	33	86.8	18	22	AA526710	Mutant chicken c-s
25	33	86.8	18	22	AA526710	Chicken pp60 c-src
26	33	86.8	18	22	AA526710	Wild-type chicken
27	33	86.8	18	22	AA526710	Mutant chicken c-s
28	33	86.8	18	22	AA526710	Mutant chicken c-s
29	33	86.8	18	22	AA526710	Amino acid sequenc
30	33	86.8	18	22	AA526710	Arabidopsis thalia
31	33	86.8	18	22	AA526710	Propionibacterium
32	33	86.8	18	22	AA526710	Human DNA/RNA bind
33	33	86.8	18	22	AA526710	Propionibacterium
34	33	86.8	18	22	AA526710	Propionibacterium
35	33	86.8	18	22	AA526710	Novel human diagno
36	33	86.8	18	22	AA526710	Human hPPG-1 prote
37	33	86.8	18	22	AA526710	Propionibacterium
38	33	86.8	18	22	AA526710	Human RTA-like G p
39	33	86.8	18	22	AA526710	Human G-protein co
40	33	86.8	18	22	AA526710	Human G-protein co
41	33	86.8	18	22	AA526710	Human G-protein co
42	33	86.8	18	22	AA526710	Human G-protein co
43	33	86.8	18	22	AA526710	Human G-protein co
44	33	86.8	18	22	AA526710	Human GPCR polyep
45	33	86.8	18	22	AA526710	Human mas proto-on

ALIGNMENTS

RESULT 1
AA526710
ID AA526710 standard; peptide; 7 AA.

AC AA526710

DT 01-NOV-1999 (first entry)

DE B. cereus T spore tight-binding peptide.

XX Bacterial spore; biopanning; phage-display library; Bacillus; vaccine;
XX pathogen detection; biological warfare agent; B. anthracis.

XX Bacillus cereus.

XX WO9936081-A1.

XX 22-JUL-1999.

XX 14-JAN-1999; 99WO-US000771.

XX 14-JAN-1998; 98US-0071411.

XX (UABR-) UAB RES FOUND.

XX Turnbough CL;

XX WPI: 1999-468943/39.

XX N-PSDB; AAX95291.

PT Identifying peptides that bind to the surface of bacterial spores by
biopanning phage-display library. useful as vaccines and diagnostic

PT agents
XX Disclosure; Page 13; 23pp; English.
XX
XX The invention provides peptides that bind to the surface of bacterial
XX spores. These peptides are identified by biopanning a phage-display
XX library with the spores. The library is incubated with spores and any
XX phage-spore complexes formed recovered by centrifuging. They are washed
XX thoroughly, then phage eluted with buffer, the eluate neutralised and
XX eluted phage amplified. The procedure is repeated, for 3 or 4 rounds of
XX panning, then individual clones purified, amplified and genomic DNA
XX extracted for determination of peptide-encoding sequences. Peptides
XX encoded by these sequences are then tested for ability to bind to target
XX spores. The peptides are used for capture and identification of
XX bacterial spores, particularly of the genus Bacillus, particularly for
XX detecting pathogens, or otherwise harmful species, in the environment
XX (e.g. air, water or food) or in clinical samples. They are also used to
XX protect against disease-causing spores (e.g. by incorporation in
XX protective masks), and in vaccines to generate a protective antibody
XX response. The peptides-specific antibodies are useful as detection
XX reagents, e.g. in enzyme-linked immunosorbent assay. The peptides are
XX species specific, i.e. they can differentiate between the possible
XX biological warfare agent B. anthracis and related species, widely present
XX in the environment and likely to give rise to false positive results.
XX Sequences AA1365-372 represent amino acid sequences of B. cereus T
XX spore tight-binding peptides.

SQ Sequence 7 AA;
Query Match 100.0%; Score 38; DB 20; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 THRLPSR 7
DB 1 THRLPSR 7

RESULT 2
AAE14550
ID AAE14550 standard; peptide; 7 AA.
XX
XX AAE14550;
XX
XX 17-MAY-2002 (first entry)
XX
XX Human alpha-synuclein aggregation inhibitor #5.
XX
XX Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;
XX Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;
XX multiple system atrophy; Hallervorden-Spatz disease; human.
XX
XX Homo sapiens.
XX
XX WO200204482-A1.
XX
XX 17-JAN-2002.
XX
XX 06-JUL-2001; 2001WO-US21379.
XX
XX 07-JUL-2000; 2000US-217319P.
XX
XX 28-MAR-2001; 2001US-279199P.
XX
XX (PANA-) PANACEA PHARM INC.
XX
XX Wolozin B, Ostretova-Golts N, Lebowitz MS;
XX
XX WPI; 2002-179695/23.
XX
XX Determination of an agent capable of inhibiting aggregation of alpha
XX synuclein useful for treating a neurodegenerative disease involves
XX PT determining aggregation of alpha synuclein in the presence of exogenous
XX iron or copper -

XX Claim 40; Page 37; 52pp; English.
XX
XX The invention relates to screening of inhibitors of alpha-synuclein
XX aggregation in the presence of exogenous iron or copper. The inhibitors
XX are magnesium and alpha-synuclein binding peptides, which are
XX useful for treating neurodegenerative disease that involves
XX the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
XX disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
XX atrophy and Hallervorden-Spatz disease. The present sequence is a
XX peptide that binds to C-terminal portion of human alpha-synuclein and
XX inhibits its aggregation.

SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 THRLPSR 7
DB 1 THRLPSR 7

RESULT 3
AAR98619
ID AAR98619 standard; Protein; 1608 AA.
XX
XX AAR98619;
XX
XX 10-DEC-1996 (first entry)
XX
XX Borna disease virus polymerase.
XX
XX Borna disease virus; BDV; G-protein; p57; nervous system disease;
XX neuro-psychiatric disease; schizophrenia; diagnosis; therapy;
XX vaccine; antibody.
XX
XX Borna disease virus strain V.
XX
XX WO9621020-A2.
XX
XX 11-JUL-1996.
XX
XX 05-JAN-1996; 96WO-US00418.
XX
XX 04-JAN-1996; 96US-0582776.
XX
XX 06-JAN-1995; 95US-0369822.
XX
XX 04-MAY-1995; 95US-0434831.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Briese T, Kliche S, Lipkin W, Schneemann A, Schneider PA;
XX Stitz L;
XX
XX WPI; 1996-333995/33.
XX
XX R-PSDB; AAT38104.
XX
XX Borna disease virus (BDV) nucleotide and protein sequences - useful
XX for the diagnosis and treatment of infection and non-BDV related
XX neuro-logic and neuro-psychiatric disease
XX
XX Claim 2; Fig 2; 186pp; English.

XX Borna disease virus (BDV) polymerase (AAR98619), or pol or p180, was
XX identified from an ORF on the virus antigenome strand (AAT38104).
XX The amino acid sequence for pol after splice modification is given
XX in AAR98605. Recombinant pol (recpol) can be expressed in transformed
XX host (partic. mammalian) cells. It is useful in assays for
XX detecting BDV infection and for diagnosing non-BDV related neurologic
XX and neuropsychiatric diseases. It may also be incorporated into
XX vaccines and used to raise anti-BDV antibodies.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-5
Perfect score: 38
Sequence: 1 THRLPSR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	7	10	US-09-901-187B-5
2	38	100.0	7	11	US-09-229-751A-72
3	33	86.8	18	9	US-09-864-761-34008
4	33	86.8	25	9	US-09-764-869-915
5	33	86.8	25	12	US-10-227-577-915
6	33	86.8	25	15	US-10-091-504-915
7	33	86.8	362	12	US-10-369-493-10705
8	32	84.2	330	12	US-10-369-493-10663
9	32	84.2	542	12	US-10-369-493-12602
10	31	81.6	71	12	US-09-864-408A-5098
11	31	81.6	321	12	US-09-920-068A-2
12	31	81.6	321	12	US-10-017-161-1050
13	31	81.6	321	12	US-10-239-421-2
14	31	81.6	321	12	US-10-258-768-1
15	31	81.6	321	12	US-10-088-726-19

16	31	81.6	321	12	US-10-305-555-4
17	31	81.6	321	12	US-10-321-807-10
18	31	81.6	321	12	US-10-237-467-8
19	31	81.6	321	12	US-09-801-94B-268
20	31	81.6	321	15	US-10-188-405-2
21	31	81.6	321	15	US-10-183-116-35
22	31	81.6	321	15	US-10-079-384-8
23	31	81.6	347	12	US-10-292-798-892
24	31	81.6	504	12	US-10-289-762-1110
25	30	78.9	43	12	US-10-029-386-27495
26	30	78.9	72	10	US-09-764-847-634
27	30	78.9	72	15	US-10-092-154-634
28	30	78.9	85	12	US-09-833-245-969
29	30	78.9	117	15	US-10-156-761-10633
30	30	78.9	186	12	US-10-158-057-223
31	30	78.9	289	12	US-10-094-749-2339
32	30	78.9	301	9	US-09-925-299-786
33	30	78.9	301	11	US-09-925-299-786
34	30	78.9	301	15	US-10-106-698-4845
35	30	78.9	389	10	US-09-738-626-4492
36	30	78.9	429	15	US-10-156-761-13110
37	30	78.9	778	15	US-10-266-388-2
38	30	78.9	797	15	US-10-266-388-4
39	30	78.9	842	11	US-09-291-417-91
40	30	78.9	846	15	US-10-303-683-22
41	30	78.9	1066	12	US-10-369-493-22668
42	30	78.9	1702	10	US-09-738-973-434
43	30	78.9	1702	10	US-09-854-133-434
44	30	78.9	1702	15	US-10-144-649A-434
45	30	78.9	6396	12	US-09-940-316B-72

ALIGNMENTS

RESULT 1

US-09-901-187B-5
; Sequence 5, Application US/0901187B
; Patent No. US20020151464A1
; GENERAL INFORMATION:
; APPLICANT: Panacea Pharmaceuticals, Inc.
; APPLICANT: Wolozin, Benjamin
; APPLICANT: Ostretova-Golts, Natalie
; APPLICANT: Lebowitz, Michael S.
; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatme
; TITLE OF INVENTION: Alpha-Synuclein Diseases
; FILE REFERENCE: PAN01/002US
; CURRENT APPLICATION NUMBER: US/09/301,187B
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/217,319
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/279,199
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-187B-5

Query Match 100.0%; Score 38; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THRLPSR 7
DB 1 THRLPSR 7

RESULT 2

US-09-229-751A-72
; Sequence 72, Application US/09229751A

Publication No. US20030044838A1
GENERAL INFORMATION:
APPLICANT: Turnbough, Charles K
TITLE OF INVENTION: PEPTIDE LIGANDS THAT BIND TO SURFACES
OF BACTERIAL CELLS
NUMBER OF SEQUENCES: 80
CORRESPONDENCE ADDRESS:
ADDRESSEE: Glenna Hendricks
STREET: P.O. Box 2509
CITY: Fairfax
STATE: VA
COUNTRY: USA
ZIP: 22031
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/229,751A
FILING DATE: 14-Jan-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hendricks, Glenna M.
REGISTRATION NUMBER: 32,535
REFERENCE/DOCKET NUMBER: turn
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 425-8405
TELEFAX: (703) 425-8406
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-09-229-751A-72
Query Match 100.0%; Score 38; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 THRLPSR 7
Db 1 THRLPSR 7
RESULT 3
US-09-864-761-34008
; Sequence 34008, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34008
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006017.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
; OTHER INFORMATION: EST_HUMAN HIT: AW997046.1, EVALUATE 2.00e-04
US-09-864-761-34008
Query Match 86.8%; Score 33; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 THRLPS 6
Db 11 THRLPS 16
RESULT 4
US-09-764-869-915
; Sequence 915, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 915
; LENGTH: 25
; TYPE: PRT